

遺伝的アルゴリズムにおける確率的揺らぎの効果

古 谷 博 史[†]

与えられた問題に遺伝的アルゴリズム(GA)を適用する際、個体数 N は直感的に決められることが多い。小さな N を選んだ場合、計算コストは軽減されるが揺らぎの効果が顕著になる。揺らぎの効果の多くは、同一試行を繰り返し平均化を行えば 0 となる。しかし、効果の中には平均化をしても有限の寄与を与えるものがある。本報告では、積型適応度関数を用いて、平均操作後の平均適応度と個体数 N の関係を調べた。数値実験の結果、個体数が少なすぎると個体数を充分多くとった計算と比較して平均適応度の増加が遅れてしまう場合があることが分かった。

Effect of Stochastic Fluctuation in Genetic Algorithms

HIROSHI FURUTANI[†]

When we apply genetic algorithms (GAs) to a given problem, we often choose the population size N intuitively. Though we can reduce the cost of calculation by choosing a small number of N , effects of random fluctuation become significant. The most part of these effects become zero after averaging repeated trials. However there are several types of effects which bring finite contribution even after averaging. In this report, we have studied the relation between the average fitness \bar{f} and N after averaging by using the multiplicative landscape. We found in numerical experiments that there is a case of slower increase of the average fitness \bar{f} with a small number of N than \bar{f} with larger N .

1. はじめに

ある与えられた問題に対して遺伝的アルゴリズム(GA)を適用する場合、集団の個体数 N は直感的に決められることが多い。難しそうな問題では N を大きく選び、そうでなければ比較的小さな N を採用する。しかし、 N を決定する客観的な基準を定めるには理論はまだ無力であり、そのための研究も少ない。

一般に個体数を少なくすると計算コストは軽減できるが、確率的揺らぎの効果は相対的に大きくなり、結果を解釈するには注意が必要となる。揺らぎの効果は、乱数を変えながら同じ計算を繰り返し、平均化を行えば消えてしまうことが多い。しかし、中には平均化の後も有限な量として観測されることがある。

我々が OneMax 問題における平均適応度と最適解頻度の世代数依存性を調べた際、両者とも個体数 N に大きく依存する場合があることが分かった¹⁾。交叉を用いずに弱い突然変異の下に GA 計算を行うと、平均適応度、最適解頻度ともに N の小さい場合は大きく決定論的モデルの結果からずれていく。二つの量は

いずれも世代ごとの増加率が小さくなり、我々にとって好ましくない結果を示す。 N を大きくしていくと決定論的モデル(N が無限に大きい場合に対応すると考えられる)とのずれは小さくなり、揺らぎの(悪い)効果は観測されなくなる。この数値実験を詳細に検討した結果、確率的揺らぎの効果は 2 次のスキーマ頻度を通じて現れることが分かった。また、1 次のスキーマ頻度は平均化の操作を行うと揺らぎの効果は消えてしまう。したがって、確率的揺らぎは高次のスキーマ頻度において有限な効果を持つことが予想された。

本報告では、論文 1) において OneMax 問題で用いた解析を、積型適応度関数の GA に適用する。また、注目する計算量としては平均適応度のみとする。積型適応度関数を用いるのは、次の理由による。この適応度を用いた決定論的モデルの下で、集団の分布は 1 次のスキーマ頻度を用いて表すことができる。したがって、高次のスキーマ頻度における揺らぎの効果がより単純な形で観測できると期待されるからである。

2. 数学モデル

確率的揺らぎの効果は、決定論的理論と数値実験とのずれに現れると考える。また、そのずれの大きさを N を変えながら調べ、揺らぎの効果の程度を評価す

[†] 京都教育大学教育学部
Faculty of Education, Kyoto University of Education

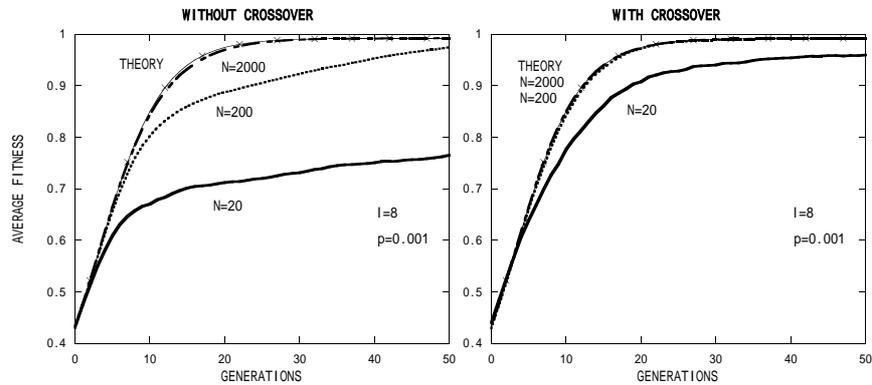


図 1 平均適応度 \bar{f} .
Fig. 1 The average fitness \bar{f} .

る．選択は適応度比例選択とし，交叉は一様交叉を用いることにする．

決定論的理論では，仮想的に個体数が無限大の集団を考え，確率的揺らぎを無視する．集団の構成要素は世代毎に入れ替わる世代的 GA を用い，進化の過程を差分方程式で記述する．集団を構成する個体は ℓ ビット固定長の 2 進ビット列で表し，これを個体の遺伝子型とよぶ．遺伝子型の種類の総数は $n = 2^\ell$ である．このビット列を非負整数の 2 進表現とみなすこともある．すなわち， i 番目の遺伝子型 B_i に対応する 2 進ビット列は， $i = \langle i_\ell, \dots, i_1 \rangle$ と表し， i を 0 から $n - 1$ の整数に対応させる．

集団内の遺伝子型分布の変化を表すため，世代 t における遺伝子型 B_i の相対頻度 $x_i(t)$ を用いる．GA の進化を記述するとき Walsh 関数 W_{ij} による Walsh 変換

$$\tilde{x}_i(t) \equiv \sum_{j=0}^{n-1} W_{ij} x_j(t) \quad (1)$$

を用いると表現が簡単になることがある²⁾．以下の式の導出も Walsh 変換を用いて得られるものが多いが，その詳細については論文 1), 2)などを参照して欲しい．

交叉や突然変異の役割を考察する上で非常に重要な概念に連鎖不平衡 (linkage disequilibrium)³⁾がある．集団が，遺伝子座 b においてビット値 m をとる相対頻度を $P[m_b]$ と表す．また，遺伝子座 b, b' においてビット値 m, m' をとる相対頻度を $P[m_b, m_{b'}]$ とする．このとき連鎖不平衡係数を次式で定義する

$$D[b, b'] = P[1_b, 1_{b'}] - P[1_b]P[1_{b'}]. \quad (2)$$

3. 積型適応度関数

積型適応度関数を次式で定義する．

$$f_i = \prod_{k=1}^{\ell} (1 - s)^{i_k}, \quad (0 < s < 1) \quad (3)$$

s は選択の強さを表すパラメータである．この適応度を最大化させる問題を考える．

3.1 連鎖平衡

論文 4) において，積型適応度における GA の進化は厳密に解けることを示した．したがって，スキーマ表現を用いた場合においても，その進化の過程を決定論の範囲で厳密に記述することができる．重要な点は，世代 $t = 0$ において連鎖平衡ならば，その後も連鎖平衡状態が維持されるということである．したがって，集団の分布はすべての 1 次スキーマの頻度が分かれば決定される．

$$x_i = \prod_{m=1}^{\ell} h(i_m), \quad (4)$$

ここで $h(i_m)$ は，ビット位置 m において定義されたビット i_m をもつ 1 次スキーマの相対頻度である．

同様に，平均適応度についても連鎖平衡状態における平均適応度を用いて

$$\bar{f}(t) = \bar{f}^{(eq)},$$

次のように表すことができる

$$\bar{f}^{(eq)} = \prod_{m=1}^{\ell} \{1 - s h(1_m)\}. \quad (5)$$

ビット当たりの適応度を次式で定義する

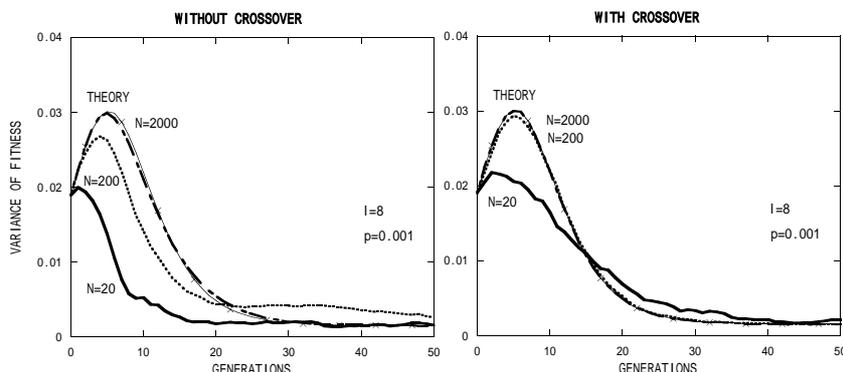


図 2 適応度の分散 $\text{VAR}(f)$.
Fig. 2 The variance of fitness $\text{VAR}(f)$.

$$g_k = 1 - s h(1_k).$$

次の節では連鎖平衡を仮定しない、一般的な平均適応度の表現を求める。

連鎖平衡状態にある集団は選択や突然変異の操作の後も連鎖平衡にあり、このとき当然すべてのスキーマも連鎖平衡状態にある。交叉は連鎖平衡を維持する操作であるから、 $t = 0$ で連鎖平衡状態にあればその後も連鎖平衡のままシステムは進化していく。

ビット当たりの適応度 g_k の選択による世代当たりの変化を

$$\Delta g_k(t) \equiv \widehat{S}g_k(t) - g_k(t),$$

とすると

$$\Delta g_k(t) = \frac{v_k(t)}{g_k(t)}, \quad (6)$$

と表される。分散 v_k は次式で定義される。

$$v_k = s^2 h(1_k) \{1 - h(1_k)\}.$$

3.2 連鎖平衡を仮定しない場合

次に、連鎖平衡を仮定しない場合の平均適応度の表現を示す。このとき平均適応度 $\bar{f}(t)$ はスキーマ頻度を用いて表現できる。

$$\begin{aligned} \bar{f}(t) = & \bar{f}^{(eq)} + s^2 \Delta^{(2)} - s^3 \Delta^{(3)} \\ & + \dots + (-s)^\ell \Delta^{(\ell)}. \end{aligned} \quad (7)$$

ここで第 m 次の項は次式で与えられる

$$\Delta^{(m)} = \sum_{k_1 < k_2} \dots \sum_{< k_m} D[k_1, k_2, \dots, k_m].$$

m 次の連鎖不平衡係数はスキーマ頻度を用いて定義される。

$$\begin{aligned} D[k_1, \dots, k_m] \\ = & h^{(m)}(1_{k_1}, \dots, 1_{k_m}) - h(1_{k_1}) \dots h(1_{k_m}). \end{aligned}$$

論文⁵⁾で示したように、OneMax 問題における適応度の分散 $\text{VAR}(f)$ は 2 つの項の和として表される。

$$\text{VAR}(f) = V_a + V_e, \quad (8)$$

$$V_a = \sum_{k=1}^{\ell} h(1_k) \{1 - h(1_k)\},$$

$$V_e = 2 \sum_{k_1 < k_2} D[k_1, k_2] = 2\Delta^{(2)}.$$

ここで V_a は additive variance, V_e は epistatic variance とよばれる。積型適応度の場合も同様に $\Delta^{(m)}$ の和として表されるが、 $\Delta^{(2)}$ から $\Delta^{(\ell)}$ の項まですべての項が現れる。しかし、その表現は複雑になるのでここでは省略する。

4. 数値実験

既に述べたように、理論的解析は決定論的仮定の下で行われた。ここでは確率的揺らぎの効果を、理論の数値実験からのずれを見ていくことにより評価することにする。数値実験は、適応度比例選択、一様交叉を用いて行った。用いたパラメータは、ビット長 $\ell = 8$ 、突然変異率 $p = 0.001$ 、選択の強さ $s = 0.2$ である。交叉率 χ は、1 (交叉あり) と 0 (交叉なし) の各々について同じ計算を行った。集団の大きさ N は、 $N = 20$ 、200、2000 の 3 通りについて同じ計算を行い、揺らぎの効果を調べた。乱数を変えながら同じ計算を 100 回繰り返し、平均化を行った。

図 1 は平均適応度の世代による変化を示したものである。左図は交叉なし ($\chi = 0$) の結果であるが、際だった N 依存性が見られる。しかも N が小さくなるに従って平均適応度の進化速度が遅れていき、 $N = 20$ ではもっとも遅れている。それに対して、 $N = 2000$

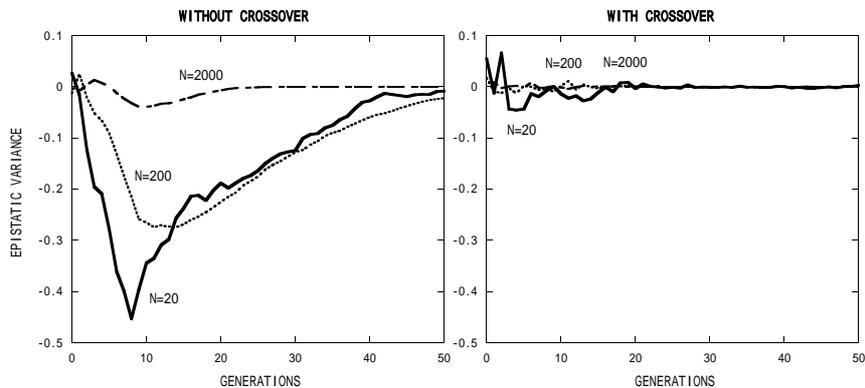


図3 2次の項, epistatic variance V_e .
Fig. 3 The second order term, the epistatic variance V_e .

の結果がもっとも良く、しかも理論値とほぼ同じ値を与えている。中間の $N = 200$ の場合は、理論値とのずれが見られ、まだ揺らぎの効果が残っているものと考えられる。 $N = 20$ と $N = 200$ の結果を見るかぎり、確率的揺らぎは進化を遅らせる方に働いているようである。

図1の右図は交叉あり ($\chi = 1$) の結果である。この図では N 依存性はあまり顕著でなく、 $N = 200$ 、 $N = 2000$ の値は理論値とほとんど同じ結果になっている。

このことをより詳しく見ていくために、図2に適応度の分散 $\text{VAR}(f)$ を示した。Fisherの定理⁶⁾によると、平均適応度の世代当たりの増加は次の式を満足する

$$\bar{f}(t+1) - \bar{f}(t) = \text{VAR}(f) / \bar{f}(t). \quad (9)$$

したがって、 $\text{VAR}(f)$ が大きいほど進化速度は速くなる。左図は交叉なし ($\chi = 0$) の結果であり、図1の左図と同様に際だった N 依存性が見られる。対照的に右図は、 $N = 20$ を除いてほとんど決定論的理論値と同じ結果になっている。この計算結果を見ると、交叉が確率的揺らぎの効果を打ち消しているように思える。

図3に $\text{VAR}(f)$ の中でもっとも大きな寄与をする高次項として、epistatic variance V_e の値を示した。右図の交叉ありの結果では、 $N = 20$ に小さな負の値が見られる他はほとんど0と考えてよい。それに対し、左図では $N = 2000$ を除いて負の値をとっていることが分かる。連鎖平衡の仮定の下では $V_e = 0$ であり、 $\text{VAR}(f)$ においてこの項からの寄与はない。しかし、揺らぎの効果からこの項が負の値をとり、結果として平均適応度の増加を遅らせている。しかし、交叉は $V_e = 0$ とする働きをし⁵⁾、結果として揺らぎの効

果を打ち消している。

5. まとめと今後の課題

確率的揺らぎは、平均化の操作を行っても大きな効果が残ることがある。積型適応度の場合は、平均適応度の増加を遅らせるという望ましくない結果を示した。しかし、交叉はこの効果を打ち消す働きをすることが分かった。しかし、どのようなメカニズムで確率的揺らぎが働いているのか、という問題は未解決であり、今後の研究が待たれる。

参考文献

- 1) 古谷博史: “選択における連鎖不平衡の効果 - OneMax 問題のスキーム解析 -,” 情報処理学会論文誌: 数理モデル化と応用, 44 掲載予定 (2003).
- 2) Furutani, H: “Schema Analysis of OneMax Problem - Evolution Equation for First Order Schemata,” K. de Jong et al. (Eds.) Foundations of Genetic Algorithms 7, Morgan Kaufmann, (2003) 19-36.
- 3) Maynard Smith, J.: Evolutionary Genetics, Oxford University Press, Oxford (1998).
- 4) H. Furutani: “Study of Evolution in Genetic Algorithms by Eigen's Theory Including Crossover Operator,” Proceedings of SEAL'00, (2000) 2696-2703.
- 5) Furutani, H: “Study of Crossover in One Max Problem by Linkage Analysis,” Proceedings of GECCO-2001, Morgan Kaufmann, (2001) 320-327.
- 6) Fisher, R.A.: The Genetical Theory of Natural selection, 2nd edition, Dover, New York (1958).