

Selection Applied within Limited Ranges in Genetic Algorithm

Tomoyuki HIROYASU* Mitsunori MIKI* and Masaki SANO**

(Received December 26, 2001)

By using Distributed Genetic Algorithm (DGA), which is one of the parallel models of Genetic Algorithm (GA), calculation time to find an optimum solution can be reduced. At the same time, DGA can find a better solution than GA with a single population. In DGA, since a population is divided into sub populations, selection is applied a smaller number of individuals. This limited range of individuals can affect the accuracy of the solutions and calculation cost. In this paper, the factors of characteristics of DGA are discussed from the point of selection operation. At first, the affection of the selection with a small number of individuals in a single population is discussed. Then, DGA and GA where the selection is performed with a small number of individuals are compared. From the discussions and the experiments, the followings are concluded. Firstly, when the selection is performed with a small number of individuals in a single population, the selection pressure becomes low. Secondly, in DGA, the diversity of the solutions is kept during the search. This maintenance of the diversity derives to find a good solution. Since the diversity can be kept when the selection pressure is low, it can be concluded that the selection with a small number of individuals in a population leads to find a good solution. Therefore, the selection with a small number of individuals is one of the factors of DGA that can find good solutions with small calculation cost.

Key words : Genetic algorithm, Distributed genetic algorithm, Selection

キーワード : 遺伝的アルゴリズム, 分散遺伝的アルゴリズム, 選択

遺伝的アルゴリズムにおける選択操作の適用範囲による解への影響

廣安知之・三木光範・佐野正樹

1. 序 論

遺伝的アルゴリズム (Genetic Algorithm : GA) は, 生物の進化を模倣した確率的な最適化アルゴリズムである¹⁾. GA では, 各探索点のある環境に棲息する生物個体とみなす. そして, 個体の集合 (母集団) に対し, より環境に適した個体を多く生き残らせる選択, 個体の一部を他の個体と交換する交叉, 個体の一部を変更する突然変異といった遺伝的操作を繰り返し適用する.

これにより, 個体集団全体が成長し, 最適解への到達が期待される. なお, GA ではこの計算の繰り返し単位を世代と呼ぶ. GA は, 目的関数の勾配情報を使用するニュートン法や最急降下法などと異なり, 各探索点の評価値のみを用いて解探索を行う. したがって連続問題にも離散問題にも適用できるという利点がある. その反面, 多数の探索点に対して評価計算を反復して行うため, 計算コストが高いという欠点がある. この

* Department of Knowledge Engineering and Computer Sciences, Doshisha University, Kyoto

Telephone:+81-774-65-6930, Fax:+81-774-65-6780, E-mail:miki@mail.doshisha.ac.jp, tomo@is.doshisha.ac.jp

** Graduate Student, Department of Knowledge Engineering and Computer Sciences, Doshisha University, Kyoto

Telephone:+81-774-65-6716, Fax:+81-774-65-6716, E-mail:msano@mikilab.doshisha.ac.jp

ため、GAの並列化に関しては多くの研究がなされてきた²⁾。

代表的な並列GA(Parallel GA:PGA)の1つに、島モデル(island model)がある^{3, 4)}。島モデルは分散遺伝的アルゴリズム(Distributed Genetic Algorithm: DGA)とも呼ばれる⁵⁾。DGAでは、個体の母集団を複数の島(サブ母集団)に分割し、島ごとにGAを適用する。そして、ある世代間隔において島間で個体の交換(移住)を行う。並列化する場合、通常はそれぞれの島をプロセッサに割り当てる。この場合、移住の際にプロセッサ間の通信が発生する。また、DGAは通常の単一母集団モデルと比較して、並列化により計算時間が短縮されるだけでなく、より適合度の高い解の発見が可能であることが報告されている^{5, 6)}。

DGAが高い性能を実現する探索メカニズムの1つとして、各島で発見された部分解が移住・交叉を経て結合し、より高品質の解が得られることが報告されている⁷⁾。これは、交叉するペアの選出が同じ島内に限られることに由来する。また、Cantú-Pazは、移住個体の選出の仕方と移住方法が解に与える影響について調査している⁸⁾。

一方、DGAでは、交叉だけでなく選択も島内ごとに行われる。言い換えれば、DGAにおける選択は島内の適合度分布にしたがう。この点も、単一母集団モデルとDGAとの大きな相違点であるため、DGAの持つ高い探索性能の一因となっている可能性がある。そこで、本論文ではこの点に関して検討を行う。

本研究では、選択の適用範囲に着目し、DGAの探索性能との関係について考察する。議論の対象とするのはルーレット選択である。本論文の前半部分は、単一母集団モデルにおいて選択の適用範囲を変化させた場合についての考察である。母集団を複数の副集団に分割して選択を適用したときの適合度分布の変化を解析する。次に、2つのテスト関数を用いた数値実験により検証を行う。また本論文の後半部分では、選択の適用範囲を限定した単一母集団モデルとDGAとの比較実験を行う。これにより、選択の適用範囲を限定することが、DGAの持つ探索性能にどのように貢献しているかについて検討する。なお、本論文では、母集団の中である操作の適用対象とする個体数をその操作の適用範囲と呼び、その個体数が母集団全体の数よりも少ない場合を適用範囲の限定と呼ぶことにする。

2. 遺伝的アルゴリズム

2.1 遺伝的アルゴリズムの基本動作

GAの基本動作をまとめると、Fig. 1のようになる。まず、初期探索点として、個体の母集団(population)をランダムに生成する。それぞれの個体(individual)には環境に対する適合度(fitness)が設定される。適合度の高い個体ほど、目的関数の評価値(evaluation value)は最適値に近くなる。そして、生成された母集団に対し、選択(selection)、交叉(crossover)、突然変異(mutation)といった遺伝的操作(genetic operator)を繰り返し適用する。遺伝的操作の繰り返し単位を、世代(generation)という。選択では適合度の高い個体ほど高い確率で生き残るようにし、交叉ではある個体の一部を他の個体と入れ替える。また、突然変異では個体の一部を変更する。

こうして世代の更新が繰り返されることによって、よりよい個体(最適解に近い個体)が増加し、やがて、最適解に到達するというのがGAの基本的な概念である。

なお、母集団内に存在する個体の数は個体数または母集団サイズ(population size)と呼ばれる。

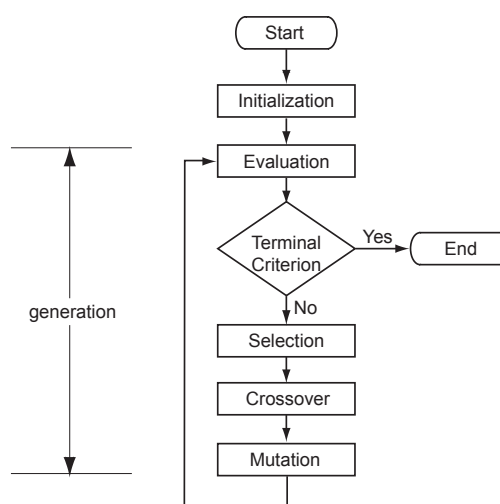


Fig. 1. GAの流れ.

2.2 選択

選択は、個体の適合度によって次世代の母集団を作成する遺伝的操作である。選択では、適合度の高い個

体ほど増殖して数が多くなり、逆に適合度の低い個体は淘汰されて死滅する (Fig. 2) . 交叉や突然変異が新しい染色体をもつ個体を生み出すのに対し、選択では生き残る個体を限定して母集団を最適解へ収束させる役割を持つ .

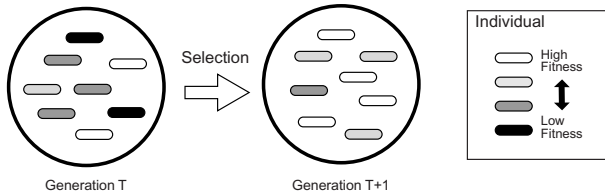


Fig. 2. 選択.

代表的な選択の1つが, Goldberg の単純 GA (Simple GA : SGA) で用いられているルーレット選択 (roulette wheel selection) である¹⁾ . ルーレット選択では, 生き残る割合が適合度の値に比例する . これについては, 次章で詳述する . この他にも, ランダムに選び出した複数の個体の中から最も適合度の高い個体を選出することを繰り返すトーナメント選択 (tournament selection) や, 適合度の順位に応じた生存確率を設定するランキング選択などがある . また, 適合度が上位の個体を無条件に次世代に残すエリート保存戦略 (elitism)⁹⁾ も選択の一部として考えられる . エリート保存戦略を用いると局所探索能力が向上する反面, 局所最適解に陥る確率も高くなる .

適合度の高い個体がきわめて多数生き残り, 適合度の低い個体がほとんど選ばれない場合, 淘汰圧 (selection pressure) が高いという . 淘汰圧は選択の強さの指標である . 探索初期において淘汰圧が高い場合, その時点で高い適合度を持つ個体のみが多く生き残り, 探索が進まないうちに解が収束してしまう初期収束 (early convergence) が起こる . 一方, 探索が進んで個体の適合度の差が小さくなると, 淘汰圧が低くなり, 選択の効果が薄れてしまうことがある .

2.3 交叉

交叉とは, 交叉率 (crossover rate) と呼ばれる一定の確率で, 個体の染色体の一部を別の個体の染色体の一部と入れ替える操作をいう . 優れた個体同士が交叉すると, 各々の個体が持つ優秀な遺伝子 (最適解を構成する部分解) が結合し, より優れた個体が生まれることが期待される .

2.4 突然変異

突然変異とは, 個体の染色体の一部を変更する操作である . 選択と交叉のみでは, 初期母集団内に存在する遺伝子の組み合わせでしか探索が行われない . そのため, 局所最適解に陥ると脱出が不可能になる . 突然変異を行うことにより, 母集団内に存在しない遺伝子を発生させ, 探索空間全域の探索が可能となる .

2.5 特徴

Goldberg によると, GA は他の最適化手法と比較して次の 4 つの特徴を持っている¹⁾ .

- 設計変数を直接操作せずに, コード化した状態で扱う .
- 1 点探索ではなく, 多点探索である .
- 目的関数の導関数やその他の付加的情報を使用せず, 関数評価値のみを用いて探索を行う .
- 決定論的規則ではなく, 確率的オペレータを用いる探索である .

GA は, 目的関数の勾配情報を使用せずに解探索を行うため, 連続問題にも離散問題にも適用できるという利点がある . また, 多峰性の関数に対しても大域的最適解を得ることができる .

しかし, GA には以下のような問題点が存在する .

- 設定すべき複数のパラメータの存在
GA においては個体数, 交叉率, 突然変異率など, 設定すべきパラメータ数が多く, またそれらの最適値は問題に依存する . そのため最適なパラメータを設定するためには, 予備実験が必要になる .
- 高い計算負荷
GA では 1 世代につき, 母集団内の個体数だけ適合度関数の評価が必要になる . そのため一点探索の最適化手法と比較して, 計算負荷が高くなる .

パラメータ設定の困難さを解消する方法としては, パラメータフリー GA (Parameter-free GA : PfGA) が提案されている^{10, 11)} .

また, 計算時間を短縮する方法として, さまざまな並列モデルが提案されている²⁾ . 2.6 節では, GA の並列モデルの 1 つである分散遺伝的アルゴリズムについて説明する .

2.6 分散遺伝的アルゴリズム

分散遺伝的アルゴリズム (Distributed Genetic Algorithm: DGA) では、母集団が分割されてサブ母集団 (subpopulation) を形成し、サブ母集団ごとに遺伝的操作を繰り返す。また、一定世代ごとにサブ母集団間で個体の交換を行う (Fig. 3, Fig. 4)。この操作を移住 (migration) という。DGA は島モデル (island model) とも呼ばれ、サブ母集団を島 (island) ともいう。

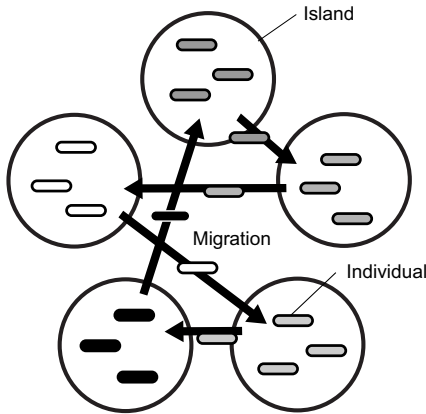


Fig. 3. 移住.

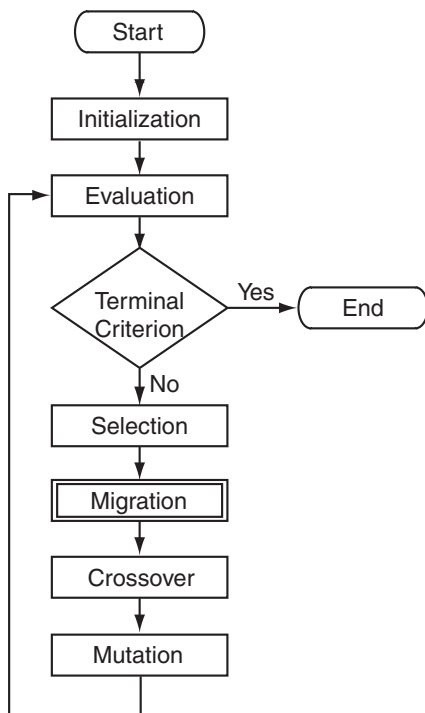


Fig. 4. 分散遺伝的アルゴリズムの流れ.

移住において、交換する個体数の、島内の個体数に

対する割合を移住率 (migration rate) といい、移住を行う世代間隔を移住間隔 (migration interval) という。

DGA には、以下のような特徴がある。

- 移住と通信

DGA は、GA の並列モデルのひとつである。DGA を並列実装する場合、通常は1つの島に1つのプロセッサを割り当てる。このとき、プロセス間通信は移住のときにのみ発生する。したがって、移住率・移住間隔によって、通信量を調節することができる。具体的には、移住率が大きいほど、また移住間隔が小さいほど、通信量は多くなる。

- 多様性の維持

DGA では、島ごとに遺伝的操作を適用する。これにより、各々の島で独立性の高い探索が進められ、母集団全体の多様性が高く保たれることが期待される⁹⁾。母集団の多様性が失われると、先に述べた初期収束などの好ましくない現象が起こる。したがって、多様性を維持する機構は DGA の大きな利点といえる。

また、DGA は単一母集団モデルと比較して、並列化により計算時間が短縮されるだけでなく、より適合度の高い解を発見することが報告されている^{5, 6)}。このような高い解探索性能を実現するメカニズムの1つとして、各島で発見された部分解が移住・交叉を経て結合し、より高品質の解が得られることが明らかとなっている⁷⁾。このメカニズムは、交叉の適用範囲が島内に限定されていることに由来する。

DGA では、交叉だけでなく、選択も島ごとに行われる。本論文では、選択の適用範囲という観点から、DGA の持つ解探索能力についての考察を行う。次章では、選択の適用範囲の変化が解探索に及ぼす影響を、理論的解析と数値実験によって検討する。

3. 選択の適用範囲の限定に関する考察

先にも述べたとおり、DGA は単一母集団モデルと比較して、並列化により計算時間が短縮されるだけでなく、より適合度の高い解を発見する。これは、DGA と単一母集団モデルとの相違点に起因すると考えられる。

DGA と単一母集団モデルとの大きな相違点の1つとして、選択の適用範囲が挙げられる。単一母集団モデルでは、全ての個体を選択の適用対象とする。これ

に対し，DGA では各島に独立に選択を適用する．個体数が 256 で島数が 8 である場合，1 世代につき 32 個体に対する選択を 8 回行うことになる．すなわち，選択の回数が増え，1 度に選択の適用対象とする個体数が少なくなるのである．

本章では，代表的な選択法であるルーレット選択を例にとり，その適用範囲の変化が及ぼす影響を理論的に解析する．また，2 つのテスト関数を用いた数値実験により，その理論の検証を行う．

3.1 選択の適用範囲の限定

ルーレット選択では，各個体が適合度に比例した形で次世代に生き残る．母集団サイズが N の場合において，適合度 f_i の個体が選ばれる確率 $P(i)$ は，式 (1) で表現される．同式において， S は母集団における適合度の合計を示す．

$$P(i) = \frac{f_i}{N} = \frac{f_i}{S} \quad (S = \sum_{j=1}^N f_j) \quad (1)$$

また，適合度 f_i の個体が選ばれる数の期待値 $E(i)$ は，式 (2) によって与えられる．式 (2) において， \bar{f} はその世代における適合度の平均を示す．

$$E(i) = NP(i) = \frac{f_i}{\bar{f}} \quad (0 \leq E(i) \leq N) \quad (2)$$

式 (2) より，ある個体が選択を経て生き残る数は，その個体の適合度を平均適合度で除した値となる．たとえば，平均適合度の 3 倍の適合度を持つ個体は，次世代には平均して 3 個体が増える．また，平均適合度と同じ適合度を持つ個体は，平均して 1 個体生き残る．

次に，母集団を d 個の副集団に分割して，副集団ごとにルーレット選択を適用する場合について検討する．副集団あたりの個体数は， $n = N/d$ である．Fig. 5 に，9 個体の母集団を 3 分割して選択を適用する様子を示す．これは，島数を d と設定した DGA に対応している．このとき，適合度 f_i の個体が選ばれる確率 $P_d(i)$ は式 (3) で表現される．また，生き残る個体数の期待値 $E_d(i)$ は式 (4) によって与えられる．これらの式において， s は，個体 i の属する副集団内の適合度の合計である．

$$P_d(i) = \frac{f_i}{n} = \frac{f_i}{s} \quad (s = \sum_{k=1}^n f_k) \quad (3)$$

$$E_d(i) = nP_d(i) = E(i) \frac{n \sum_{j=1}^N f_j}{N \sum_{k=1}^n f_k} = E(i) \frac{S}{d \cdot s} \quad (4)$$

$(0 \leq E_d(i) \leq n)$

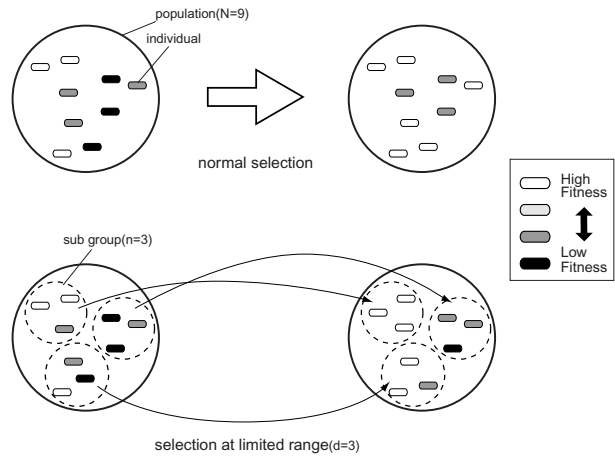


Fig. 5. 選択の適用範囲の限定.

通常のルーレット選択では，ある個体が選択の後に残る数は最大で N である．これに対し，選択の適用範囲を限定した場合にはその最大値が n となる．適合度の高い個体がこの影響を強く受ける．たとえば，ある個体の適合度が平均適合度の 5 倍であるとき，通常のルーレット選択ではその個体が生き残る数の期待値は 5 である．一方，選択の適用範囲を 3 個体ずつに分割した場合，どれだけ適合度が高い個体も 3 個体より多く増殖することはない．これにより，適合度が突出した個体が母集団の大部分を淘汰する現象が回避されることになる．

式 (4) において， d と S は個体によって変化しないため，定数と見なせる．個体によって変わるのは， s である． s が S/d と等しい場合には， $E_d(i)$ と $E(i)$ は等しくなる．また， f_i の値が大きいたときには s の値も大きくなる可能性が高い．このときには， $E_d(i)$ の値が $E(i)$ よりも小さくなる．つまり，ルーレット選択の適用範囲を限定すると，適合度の突出した個体はその増殖が抑えられるため，淘汰圧が低くなる．

本章では、以上の議論を検証するため、数値実験を行う。より具体的には、淘汰圧が高いほど良い解に到達する可能性が高いと予想される状況と、淘汰圧が高いと解の精度が悪化すると予想される状況とについて、選択の適用範囲の変化が解探索に与える影響を調査する。次節では、数値実験に使用する2つのテスト関数について説明する。

3.2 対象問題

本章の数値実験では、Table 1 に示す2つのテスト関数を導入する。いずれも最大化問題である。同表において、 L は染色体長であり、 s_i は染色体のそれぞれのビットを指す。また、 F_b においては必ず4の倍数であることとする。2章で述べたように、 s_i は $\{0, 1\}$ の値をとる。

F_a では、遺伝子の配列において'1'であるビット(以下、ビット'1')の個数を Num_{one} として、 Num_{one} の5乗がその関数値となる。最適値(最大値)は染色体長 L の5乗である。Fig. 6 にその外形を示す。同図において、横軸は Num_{one} を示し、縦軸は適合度値を示す。Fig. 6 より、 Num_{one} が設計変数である関数としてみた場合、 F_a は局所的最適解の無い単調増加の関数となる。

F_b では、 Num_{one} が4の倍数であるときには Num_{one}^2 を関数値として採用し、それ以外は $(Num_{one}/2)^2$ を関数値とする。 L を4の倍数と決めた場合には、最適値は L の2乗となる。Fig. 7 にその外形を示す。同図より、 Num_{one} が設計変数である関数としてみた場合、 F_b には局所的最適解が存在する。たとえば、 Num_{one} が16のとき、ビット'1'が1つ増減するとその評価値が下がる。

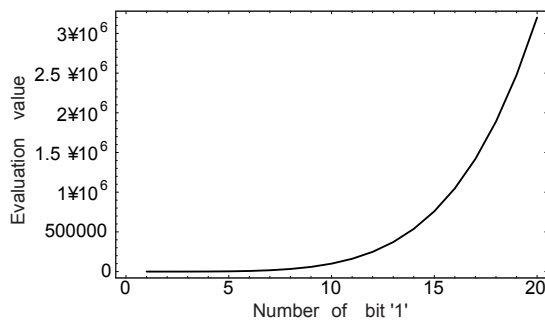


Fig. 6. F_a .

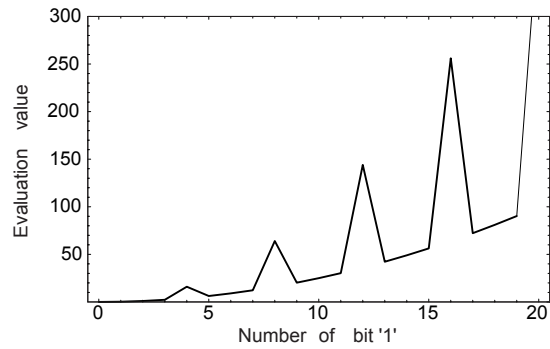


Fig. 7. F_b .

3.3 実験内容

本章では、2つの状況について、通常の単一母集団モデルと母集団を8分割して選択を適用したものの比較を行う。1つ目は、淘汰圧が高いほど良い解に到達する可能性が高いと予想される状況であり、2つ目は、逆に淘汰圧が高いと解の精度が悪化すると予想される状況である。以下に、その詳細を記す。

1つ目の実験では、交叉率を0、突然変異率を $1/L$ と設定し、 F_a を最大化する。個体の遺伝子を変更する操作は突然変異のみとなる。個体間の情報交換が無く、適合度値はビット'1'の数にのみ依存するため、母集団の多様性を維持する必要は無い。たとえば、染色体"00110"、"11101"をそれぞれもつ2つの個体を1ビット反転するとき、前者の適合度は1または 3^5 になるのに対し、後者のそれは 5^5 または 3^5 になる。前者は後者よりもビット'0'の数が多いので、適合度が上がりやすい。しかし、前者の個体もビット'1'が多くなると適合度が上がりにくくなるため、結局はその時点で適合度の高い個体ができるだけ多いほうが良いことになる。以上の議論より、この状況では、選択において淘汰圧が高い方が望ましいと予想される。

他方の実験では、交叉率を0.8、突然変異率を0とし、 F_b を最大化する。この場合、初期母集団内に存在する遺伝子の組み合わせでしか探索が行われないので、できるだけ多くの組み合わせで交叉を行うために、母集団の多様性を維持する必要がある。また、局所的最適解に陥った個体が選択において急激に増殖すると、それ以上探索が進まなくなる恐れがある。以上より、この状況では、選択において高い淘汰圧がかかると得られる解の精度が悪化すると予想される。

その他のパラメータとしては、個体数を32、染色体長を20とし、交叉法には1点交叉を、選択にはルー

Table 1. テスト関数 ($m = 1, 2, \dots$).

Function equation	Optimum
$F_a = \left(\sum_{i=1}^L s_i\right)^5$	L^5
$F_b = \left(\frac{\sum_{i=1}^L s_i}{\delta}\right)^2 \quad \delta = \begin{cases} 1 & (\sum_{i=1}^L s_i = 4m) \\ 2 & (\sum_{i=1}^L s_i \neq 4m) \end{cases}$	$L^2 \quad (L = 4m)$

レット選択を用いた．母集団を 8 分割して選択を適用する場合，各副集団には 4 個体が含まれることになる．また，各世代 1 個体をエリートとして保存する．

3.4 結果と考察

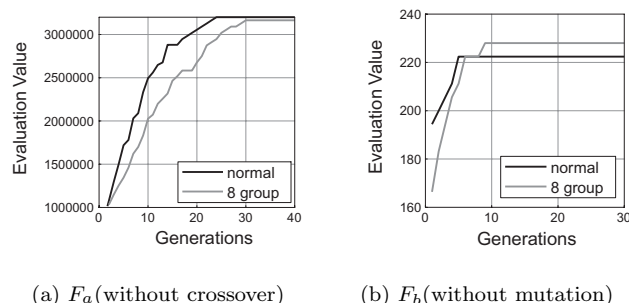
Fig. 8 は，横軸に世代をとり，最良個体の目的関数値の推移を示したものである．目的関数値が大きいほど，より良好な解を得ていることを示す．同図において，(a) は F_a に対する実験の結果であり，(b) は F_b に対する実験の結果である．また，通常の単一母集団モデルを normal，母集団を 8 分割して選択を適用したものを 8group と表現する．以下に，それぞれについての考察を行う．

(a) F_a ，交叉無し この実験では，目的関数値の最大は 3200000 である．normal では，解の収束が速く，30 世代までに最適解が得られている．これに対し，8group では，40 世代を経ても最適解が得られていない．3.1 節の議論より，8group は normal よりも淘汰圧が低い．よって，normal の性能が勝ることは，3.3 節の予想に反しない．

(b) F_b ，突然変異無し 探索の初期においては normal の方が解の収束が速いが，探索が進むと 8group の方が良い解を発見している．3.1 節の議論に照らし合わせれば，8group では淘汰圧が抑えられるために緩やかに解が収束し，最終的に normal よりも良好な解に到達すると説明できる．よって，3.1 節の内容に反しない結果であるといえる．

以上の結果より，淘汰圧が高い方が望ましいと予想した状況では normal の探索性能が勝り，逆にそれが望ましくないと予想した状況では 8group が勝っている．よって，3.1 節の議論に沿った結果であるといえる．

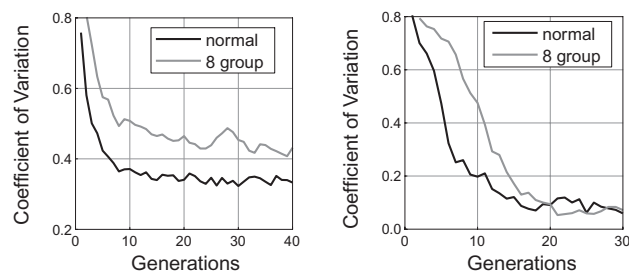
次に，目的関数値の変動係数の推移を，Fig. 9 に示す．変動係数は標準偏差を平均で割ったものであるので，この値が大きいほど母集団の多様性が保たれてい



(a) F_a (without crossover) (b) F_b (without mutation)

Fig. 8. 目的関数値の変動係数の推移.

るといえる．(a) は F_a に対する実験の結果であり，(b) は F_b に対する実験の結果である．(a) と (b) いずれの場合においても，探索初期においては 8group の方が変動係数が大きい．淘汰圧が低いと，適合度の高い個体が急速に母集団に広がることが抑制され，母集団の多様性が保たれることが期待される．よって，この結果は 3.1 節の議論と矛盾しない．また，(b) において探索が進んだ後に normal，8group 両方の変動係数が小さくなるのは，突然変異を行わないために新しい遺伝子の組み合わせが生まれにくくなっているためと考えられる．



(a) F_a (without crossover) (b) F_b (without mutation)

Fig. 9. 最良個体の目的関数値の推移.

本節の実験結果は、全て 3.1 節で述べた事柄にしたがっているといえる。すなわち、母集団を複数の副集団に分割してルーレット選択を適用することは淘汰圧を抑えることにつながり、これによって多様性を維持する効果があると考えられる。

4. 分散遺伝的アルゴリズムと選択の適用範囲

前章では、母集団を複数の副集団に分割してルーレット選択を適用することが、解の挙動にどのように影響するのかについて述べた。単一母集団モデルにおいて、母集団を d 分割して選択を行う場合、その選択の適用範囲は d 島の DGA と一致する。そこで本章では、両者の比較実験を行うことにより、選択の適用範囲を限定することが DGA での解探索能力に及ぼす影響について検討する。なお、用いる選択はルーレット選択である。

4.1 対象問題

本章の数値実験における対象問題は、Table 2 に示す 4 つのテスト関数の最小化である。いずれも 30 次元のものを用いる。Rastrigin 関数・Griewank 関数・Ridge 関数の最適値は 0 である。Schwefel 関数の最適値は 0 ではないが、本論文では最適値が 0 となるように定数を加えている。

GA の数値実験で用いるテスト関数では、設計変数間の依存関係が重要視される¹²⁾。設計変数間に依存関係の無い m 次元の関数は、 m 個の関数の線形結合として表現できる。たとえば、3 次元の場合は式 (5) のようになる。同式において、 $F(x, y, z)$ は 3 つの設計変数 x, y, z を持つ関数である。設計変数間に依存関係が無い $F(x, y, z)$ は、 x の関数 $S_1(x)$ と y の関数 $S_2(y)$ と z の関数 $S_3(z)$ の線形結合として表現される。この場合、 $F(x, y, z)$ の最適化は、3 つの関数 $S_1(x) \cdot S_2(y) \cdot S_3(z)$ の最適化に分離できる。

$$F(x, y, z) = S_1(x) + S_2(y) + S_3(z) \quad (5)$$

このように、設計変数間の依存関係の有無によって問題の性質が異なる。よって、設計変数間に依存関係の有る問題と無い問題の両方について実験を行うことにより、より一般性の高い議論が可能となる。Table 2 に示す関数のうち、Rastrigin 関数と Schwefel 関数では設計変数間に依存関係があり、Ridge 関数と Schwefel 関数では設計変数間に依存関係が無い。Fig. 10 に各関数の 2 変数の場合の外形を示す。

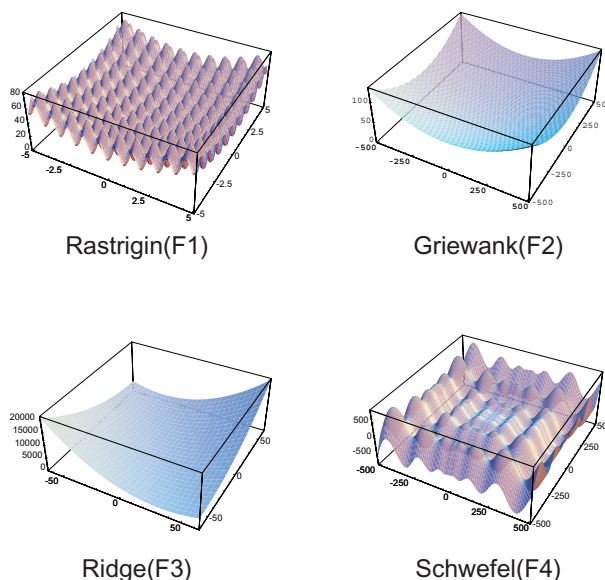


Fig. 10. テスト関数.

4.2 実験内容

本章の数値実験では、単一母集団モデルにおいて選択の適用範囲を限定したものと、DGA とを比較する。母集団を d 分割して選択を適用した単一母集団モデルでは、その選択の適用範囲は d 島の DGA と一致する。たとえば、母集団サイズが 9 のとき、母集団を 3 分割して選択を適用したものと 3 島の DGA との対応は、Fig. 11 のようになる。本論文では、この d を対応させて両者を比較することにより、DGA における選択の適用範囲の限定の効果について議論する。

より具体的な数値実験の方法としては、DGA の島数を 1,2,4,8,16,32 としたものと、それに対応して母集団を分割して選択を適用したものとを比較する。その他のパラメータについては、母集団サイズを 128、交叉率を 0.8、突然変異率を $1/L$ (L : 染色体長) と設定した。DGA の移住間隔と移住率は、それぞれ 5 世代、0.3 とした。また、各世代 1 個体をエリートとして保存する。DGA では、各島につきエリートが 1 個体である。

4.3 結果と考察

4.3.1 最良個体の目的関数値の推移

Fig. 12 と Fig. 13 に、最良個体の目的関数値の推移を示す。これらの図において、横軸は個体の評価回数である。4 つのテスト関数はいずれも最小化問題であるため、目的関数値が小さいほど、良い解を得ていることを示す。この 2 つの図のうち、Fig. 12 は DGA

Table 2. テスト関数.

	Function equation	Name	Constraint
F1	$f_1 = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i))$	Rastrigin	$-5.12 \leq x_i < 5.12$
F2	$f_2 = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \left(\cos\left(\frac{x_i}{\sqrt{i}}\right) \right)$	Griewank	$-512 \leq x_i < 512$
F3	$f_3 = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	Ridge	$-64 \leq x_i < 64$
F4	$f_4 = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	Schwefel	$-512 \leq x_i < 512$

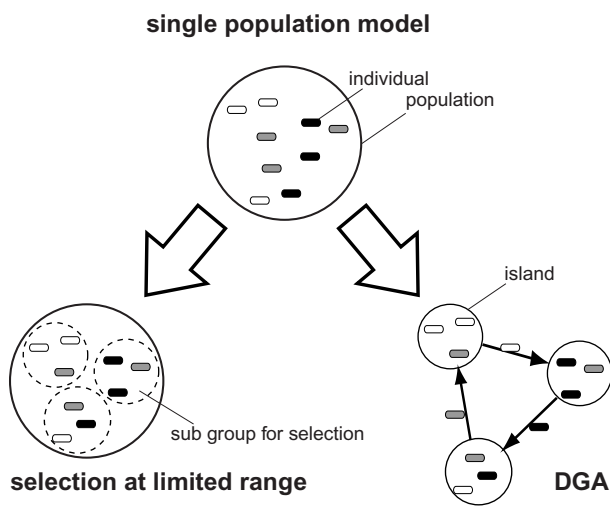


Fig. 11. 選択の適用範囲を限定した単一母集団モデルとDGA.

の実験結果である．同図において， d island は d 島のDGAを示す．なお，1 island は通常の単一母集団モデルを示す．また，Fig. 13 は，選択の適用範囲を変化させた単一母集団モデルの実験結果である．同図において，normal は通常の単一母集団モデルを示し，選択の適用範囲を d 分割したものは d group と表現する．ここで，normal は，Fig. 12 の 1 island と同一である．

Griewank 関数と Ridge 関数に着目する．Fig. 13 より，この2つの関数に対しては，単一母集団モデルにおいて選択の適用範囲を多数に分割するほど良い解を得る傾向がある．特に Ridge 関数はその傾向が強い．また Fig. 12 より，これらの関数に対しては，DGA においては島数を多くするほど良い解を得ている．よって，Griewank 関数と Ridge 関数には，選択の適用範囲を限定することによって良い解を得るという性質が

あり，このことがDGAが高い探索性能を示す要因の1つになっているといえる．

次に，Rastrigin 関数と Schwefel 関数に着目する．Fig. 13 より，この2つの関数に対しては，単一母集団モデルにおいて，選択の適用範囲の分割数と探索性能との間に有意な関係は見られない．しかし，Fig. 12 より，Rastrigin 関数ではDGAが単一母集団モデルと比較して高い解探索性能を示している．Schwefel 関数では，DGAにおいて島数を増やすほど速く良い解を発見する傾向が見られる．よって，選択の適用範囲の限定のみがDGAの解探索性能を向上させているのではないといえる．

4.3.2 目的関数評価値の平均と変動係数

Fig. 14 と Fig. 15 に，目的関数値の平均の推移を示す．この値が小さいほど，より母集団全体が最適解付近に集中していることを示す．図中の横軸や凡例の意味は，Fig. 12・Fig. 13 と同様である．なお，DGA については，各島の平均ではなく母集団全体の平均値を求めた．

Fig. 14 より，DGA においては，島数が多くなるほど目的関数値の平均が小さくなる傾向がある．また，Fig. 15 においては，Schwefel 関数以外に対して，選択の適用範囲を多数分割した方が目的関数値の平均が小さくなっている．したがって，選択の適用範囲の限定は，DGA における平均目的関数値の挙動を特徴付ける1つの要因になっているといえる．

次に，Fig. 16 と Fig. 17 に，目的関数評価値の変動係数の推移を示す．この値が大きいほど，目的関数値のばらつきが大きいので，母集団内の多様性が維持されているといえる．図中の横軸や凡例の意味は，Fig. 12・Fig. 13 と同様である．なお，DGA については，各島ではなく母集団全体の目的関数値の変動係数を求

めた。

この2つの図より、Schwefel関数以外では、選択の適用範囲を多数分割するほど、あるいは島数が多いほど、変動係数が大きくなる傾向が見られる。

以上より、本項の実験結果から次のことがいえる。DGAでは、島数が増加するにしたがって、母集団全体の目的関数値を最適値に近づける効果が強くなる。さらに、平均目的関数値が最適値に近づいた場合でも、母集団の多様性は維持される。このことは、選択の適用範囲を限定した単一母集団モデルと共通する。すなわち、選択を適用する副集団の数が多いほど、母集団全体の目的関数値は最適値に近づくが、その場合でも多様性は維持される。したがって、DGAには、多様性を維持しながら母集団全体を最適値に近づけるといった特性があり、選択の適用範囲の限定がその1つの要因になっているといえる。

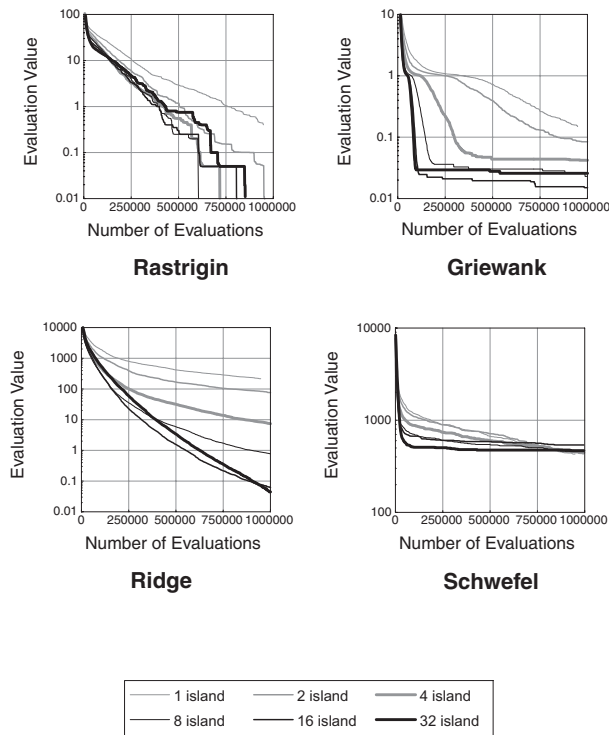


Fig. 12. DGAにおける最良個体の目的関数値の推移.

5. 結論

本論文では、選択の適用範囲に着目し、DGAの探索性能との関係について検討した。議論の対象としたのはルーレット選択である。

3章では、単一母集団モデルにおいて選択の適用範囲を変化させた場合についての考察を行った。その結

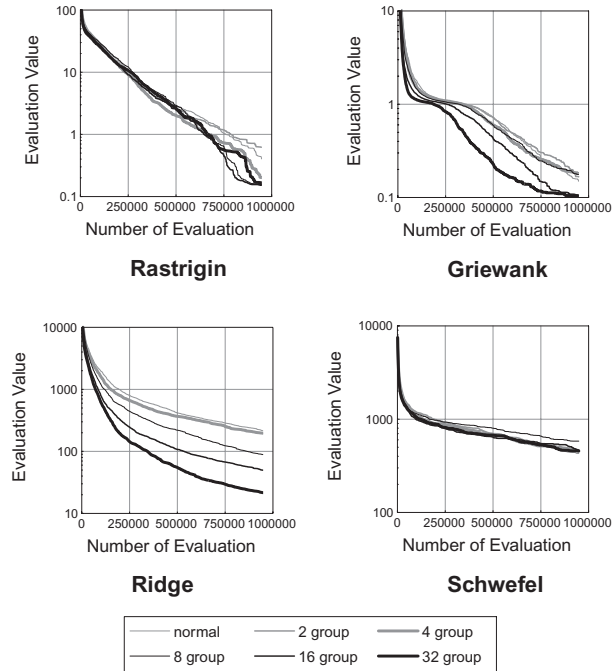


Fig. 13. 選択の適用範囲を限定した単一母集団モデルにおける最良個体の目的関数値の推移.

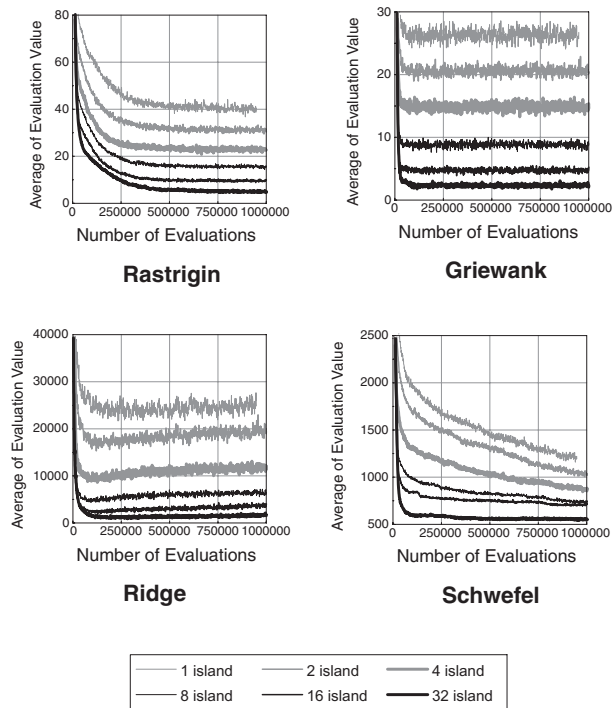


Fig. 14. DGAにおける目的関数値の平均の推移.

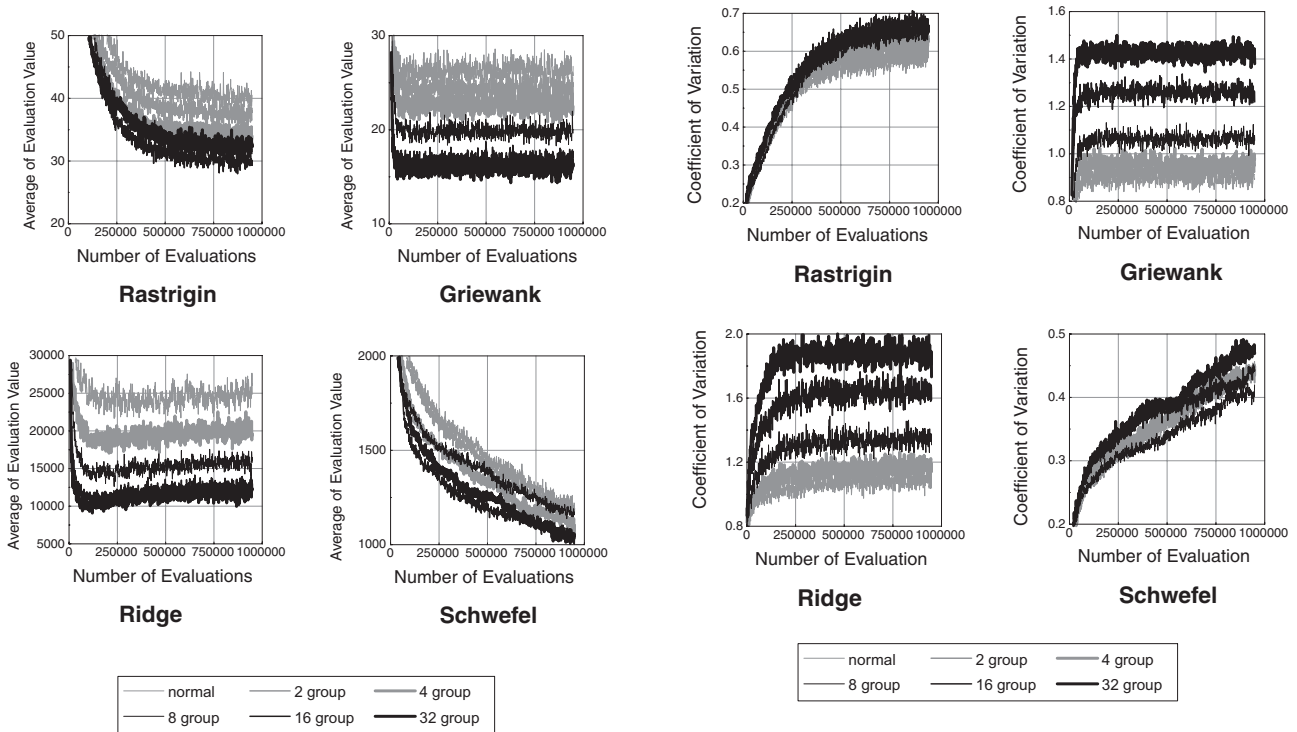


Fig. 15. 選択の適用範囲を限定した単一母集団モデルにおける目的関数値の平均の推移.

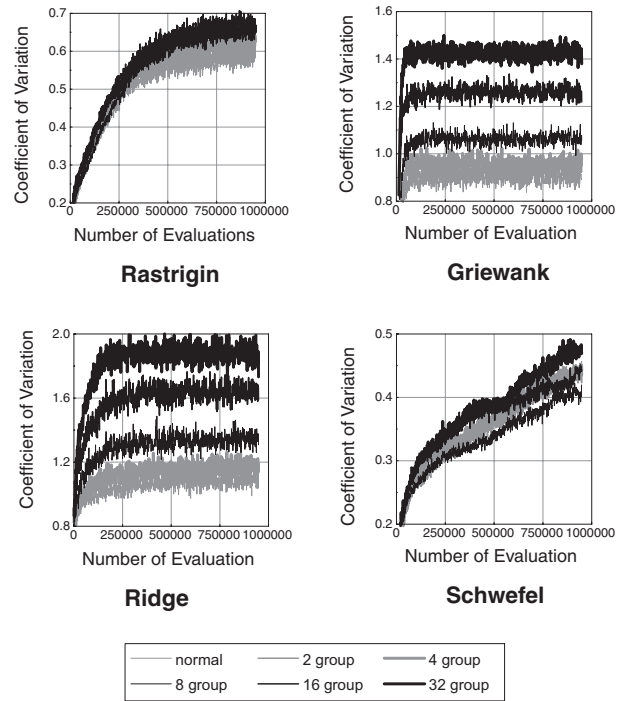


Fig. 17. 選択の適用範囲を限定した単一母集団モデルにおける目的関数値の変動係数の推移.

果、選択の適用範囲を限定することは、淘汰圧を抑える効果があることが明らかとなった。

4章では、一般的な4つのテスト関数を用いて、選択の適用範囲を限定した単一母集団モデルとDGAとの比較実験を行った。その結果は以下の通りである。

- 最良個体の目的関数値の推移について

Griewank 関数と Ridge 関数に対しては、単一母集団モデルにおいて選択の適用範囲を多数に分割するほど良好な解を得る。また、これらの関数に対しては、DGAにおいて島数を多くするほど探索性能が向上する。よって、Griewank 関数と Ridge 関数には、淘汰圧を抑えることによって良好な解を得るといった性質があり、このことが島数の多いDGAにおいて高い解探索性能を示す1つの要因となっていると考えられる。

Rastrigin 関数と Schwefel 関数については、選択の適用対象とする副集団数と探索性能との間に有意な関係は見られない。しかし、これらの関数にDGAを適用すると、単一母集団モデルと比較して高い解探索性能を示す。よって、選択の適用範囲の限定のみがDGAの解探索性能を向上させているのではないといえる。

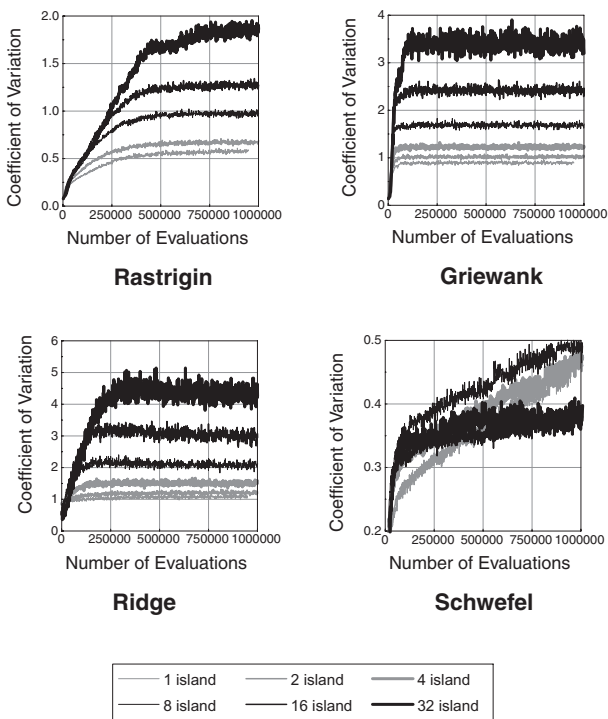


Fig. 16. DGAにおける目的関数値の変動係数の推移.

● 目的関数値の平均と変動係数の推移について

4章で用いた4つのテスト関数のうち, Schwefel関数以外で以下のことが確認された. DGAでは, 島数が増加するほど平均目的関数値が最適値に近づき, 目的関数値の変動係数は大きくなる. 単一母集団モデルにおいても, 選択を適用する副集団の数が多いほど平均目的関数値は最適値に近づき, 目的関数値の変動係数は大きくなる.

よってDGAには, 多様性を維持しながら母集団全体が最適値に近づく特性があり, 選択の適用範囲の限定がその1つの要因になっているといえる.

6. 今後の課題

今後の課題として以下を挙げる.

他の選択手法についての調査

本研究では, 選択手法としてルーレット選択のみを解析・実験の対象とした. よって, 今後の展開として, 他の代表的な選択手法であるトーナメント選択やランキング選択を対象とした議論が考えられる. また, 淘汰圧を調節する方法として, スケーリングがある. スケーリングを用いることにより, 選択の適用範囲と淘汰圧との関係について, さらに調査を進めることができると予想される.

エリート保存

3章の解析はエリート保存を想定していない. 本論文の数値実験ではエリート保存を用いているが, それが解探索に与える影響については考察していない. エリート保存の方法によっても解の挙動が変化すると予測されるので, この点について調査が必要である.

参 考 文 献

- 1) D.E.Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*. Addison-Wesley, 1989.
- 2) E. Cantú-Paz, A survey of parallel genetic algorithms. *Calculateurs Paralleles*, Vol. 10, No. 2, 1998.
- 3) C. C. Petty and M. R. Leuze, A theoretical investigation of a parallel genetic algorithm. *Proc. 3rd International Conference on Genetic Algorithms*, pp. 398-399, 1989.
- 4) H. Mülenbein and J. Born M. Schomisch, The parallel genetic algorithm as function optimizer. *Proc. 4th International Conference on Genetic Algorithms*, pp. 271-278, 1991.
- 5) R. Tanese, Distributed genetic algorithms. *Proc. 3rd International Conference on Genetic Algorithms*, pp. 434-439, 1989.
- 6) T. C. Belding, The distributed genetic algorithm revisited. *Proc. 6th International Conference on Genetic Algorithms*, pp. 114-121, 1995.
- 7) 三木光範, 廣安知之, 金子美華, 分散母集団遺伝的アルゴリズムにおける解探索能力. 人工知能学会全国大会, 1999.
- 8) E. Cantú-Paz, Migration policies, selection pressure, and parallel evolutionary algorithms. *IlligAL Report*, No. 99015, Jun 1999.
- 9) 三宮信夫, 喜多一, 玉置久, 岩本貴司, 遺伝的アルゴリズムと最適化. システム制御情報ライブラリー 17. 朝倉書店, 1998.
- 10) 足立進, 澤井秀文, 並列分散パラメータフリー遺伝的アルゴリズムにおける移民選択法の効果. 電子情報通信学会論文誌, Vol. J83-D-I, No. 8, pp. 834-843, 2000.
- 11) H. Sawai and S. Adachi, Genetic algorithm inspired by gene duplication. *Congress on Evolutionary Computation*, Vol. 1, pp. 480-487, 1999.
- 12) D. Whitley, K. Mathias, S. Rana and J. Dzubera, Building better test functions. *International Conference on Genetic Algorithms*, 1995.

出典：

同志社大学理工学研究報告 Vol.43, No.1

pp. 29-40

(2002年4月)

問い合わせ先：

同志社大学工学部/ 同志社大学大学院工学研究科

知的システムデザイン研究室

(<http://mikilab.doshisha.ac.jp>)