

単峰性関数当てはめによるGA高速化の評価

印具, 毅雄
九州芸術工科大学大学院

高木, 英行
九州芸術工科大学芸術情報設計学科

<https://hdl.handle.net/2324/4481611>

出版情報 : pp.117-120, 1999-06-02. 日本知能情報ファジィ学会
バージョン :
権利関係 :



単峰性関数当てはめによる GA 高速化の評価

Evaluation of Accelerating a GA Convergence by Fitting a Single-Peak Function

印具 毅雄[†]高木 英行[‡]Takeo INGU[†] Hideyuki TAKAGI[‡]九州芸術工科大学[†] 大学院, [‡] 芸術情報設計学科Kyushu Institute of Design[†] Graduate School, [‡] Dept. of Art and Information Design

Abstract: We propose to approximate a GA search surface by fitting a single-peak function to accelerate GA convergence and evaluate its performance using benchmark tests. As most rough approximation of a finite searching surface that has one global optimum is anticipated to be a single-peak curve that has only one peak, the vertex of the approximated single-peak function is expected to locate near the global optimum. We show three methods to select data for the fitting, use a quadratic function as the single-peak function, and evaluate the proposed idea using seven benchmark functions. The experiment has shown that the proposed method surely accelerates GA convergence per generation.

1 はじめに

多くのアプリケーションで GA の高速化が必要とされている。たとえばファジイコントローラ設計の場合、1 個体の評価値を求めるのに長い時間がかかるため GA 演算の計算コストよりもむしろ早い世代での収束が求められている。このような必要性は、ヒューマンインタフェースを伴う GA の場合では通常の GA よりも顕著である [4]。

著者らは図 1 のように、探索空間を単峰性関数で近似し、GA の収束を高速化する手法を研究している。

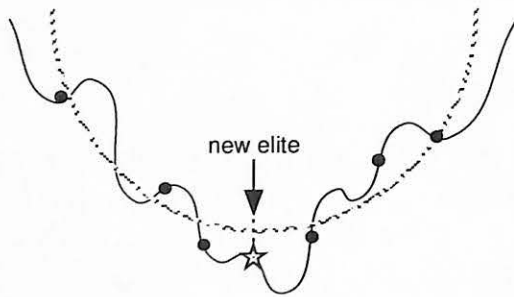


図 1: 探索空間に単峰性関数当てはめの様子と、生成された新しいエリート個体。

文献 [3] では単峰性関数として二次形式の関数を用いている。ここで、二次形式の関数の次元数は遺伝子の数と同じである。たとえば、染色体に n パラメータをエンコードした場合、つまり n 次元探索空間の場合、探索空間を近似する二次形式の関数は以下式 (1) のように表される。

$$f(x_1, x_2, \dots, x_n) \approx \sum_{i=1}^n a_i (x_i - b_i)^2 + c \quad (1)$$

$(x_1, x_2, \dots, x_n)$ は GA のパラメータを表しており、 $f(\mathbf{x})$ は個体に与えられた評価値を示す。

求めた二次関数の極小値は $x_i = b_i$, ($i = 1, \dots, n$) のとき与えられ、新しいエリート個体となる。したがって a_i , c は本手法にとって重要な値ではない。新しいエリート個体は、現世代の最低の評価値を持つ個体と比較され、評価値が上回れば入れ替えられる。

本論文では、文献 [3] で提案されている手法を基本として、他の有効な手法を提案し、評価することを目的としている。また、単峰性関数として文献 [3] と同様に二次形式の関数を用いている。

二次形式の関数は LS 法 (the least-squares method) によって求められ、使用するデータとして、過去を含めたすべての個体データ、評価値基準の優良 n データを使用する方法 (評価値基準データ選択法) を提案していた。本論文では、GA 探索空間への二次関数近似方法として新たに SubGA 法を提案し、使用データの選択方法として、エリート距離基準の近傍 n データを使用する方法 (エリート距離基準データ選択法) を提案する。これらの手法を LS 法、評価値基準データ選択法と比較、評価する。

2 GA 探索空間の二次関数近似

2.1 SubGA 法による二次関数当てはめ

式 (1) の (a_i, b_i, c) を求めるということは、最適化問題を解くことに等しいため、GA でこれらの係数を決定する。高速化したい本来の GA と区別するために、この GA を SubGA と呼ぶことにする。LS 法は長い計算時間の後、正確な (a_i, b_i, c) を得る。SubGA も時間をかければ同様に精度

を上げることができるが、本来が近似であるから正確な二次形式の関数を必要としない。もしラフなエリート個体を生成する SubGA 法と LS 法に性能の差がないならば、計算コストの少ない SubGA 法は LS 法よりも優れている。メイン GA の探索状況によって適応的に SubGA の世代数を変化させれば、さらに計算コストを軽減することができるだろう。

SubGA によって適当な二次形式の関数が得られたなら、 b_i が新しいエリート個体となり、現在の評価値最低の個体より評価値がよければ入れ替えられる。

2.2 エリート距離基準による近傍 n データ選択 (エリート距離基準データ選択法)

式 (1) の (a_i, b_i, c) を決定するために過去の探索点 (個体) を使用する必要がある。もっとも単純な方法は、すべての過去の個体を使用する方法である。この方法は探索空間全体に当てはめ、新しいエリート個体を得ることができるが、計算コストが世代数と個体数を乗じたものになる。計算コストに加えて、探索空間全体の形状と二次形式の関数との差異にも注意を払う必要がある。最適解近傍では二次形式に近いといわれており、二次形式の関数近似は、探索空間全体への近似よりも良いことが多いと考えられる。我々の関心は探索空間全体の形状ではなく、最適解である。したがって、よりよいデータを用いた二次形式の関数あてはめは、計算コストだけでなく、優れたエリート個体の生成も期待できる。

上述した2点に着目して、文献 [3] では評価値基準データ選択法を提案していた。しかし、評価値基準データ選択法では、局所解近傍にある個体と最適解近傍にある個体の評価値が近く、位置が離れている場合、二次形式の関数の極小値はそれらの中間地点に位置する。図 2 (a) の探索空間がその例である。このような場合でも回避できる方法として、エリート距離基準データ選択法を提案する。

本手法は、現在と過去の世代の個体から現在のエリート個体と距離が近い n 個体を選び出す方法である。最適解近傍を集中的に探索するのが本提案手法の特徴であり、さらに計算時間は評価値基準データ選択法と同様に軽減される。図 2 (b) はこれらの手法によって選択されている様子を示したものである。

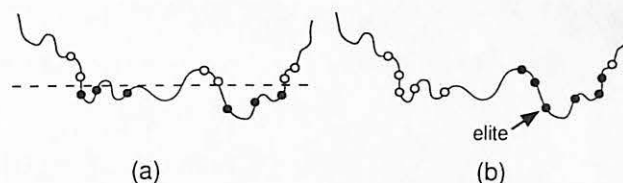


図 2: 2つのデータ選択方式。黒点、白点はそれぞれ選択された個体、選択されなかった個体を示す。(a) 評価値基準データ選択法では現在のエリートの評価値に近い n 個体が選択されており、(b) エリート距離基準データ選択法では現在のエリートから距離の近い n 個体が選択される。

3 評価実験

文献 [3] および本論文で提案した GA 高速化手法を DeJong の 5 関数 [1], Schaffer の 2 関数 [2] を用いて評価した。二次形式の関数を LS 法によって決定する提案手法、SubGA 法によって決定する提案手法と従来の GA による方法の 3 つのグループについて評価を行った。前者の 2 つの提案手法については 2.2 で述べたデータ選択方法をもとに 3 つに分類されている。以上の条件を表 1 に示す。

GA の収束曲線 (図 3) は、異なる初期乱数を 30 種類用いて収束曲線を平均したものである。GA の実験条件を表 2 に示す。

4 考察

我々は提案手法が従来の GA と比べて有意に良くなっていることを調べるために、統計検定を行った。標本が正規分布に従うと仮定して、母平均の差の検定を各世代毎に行った。今回の実験は 30 回行っているの、世代では各比較手法毎に 30 のサンプルがある。この各平均が図 3 に示したグラフである。この 30 サンプルが各世代毎に正規分布に従うと考えれば、平均の差の検定を行うことができる。この検定を各世代毎で行う。

以上の評価実験の結果、7 つのベンチマーク関数すべてにおいて、危険率 1% または 5% で、2 つの提案手法のうち少なくとも 1 つは、従来の GA より有意に収束が早くなっていることが示された。すべての実験結果から、3 つのデータ選択方法間に有意な差はみられなかった。計算コストの面から評価値基準データ選択法が優れていることがいえよう。

本提案手法は初期の世代で特に有効に働いて

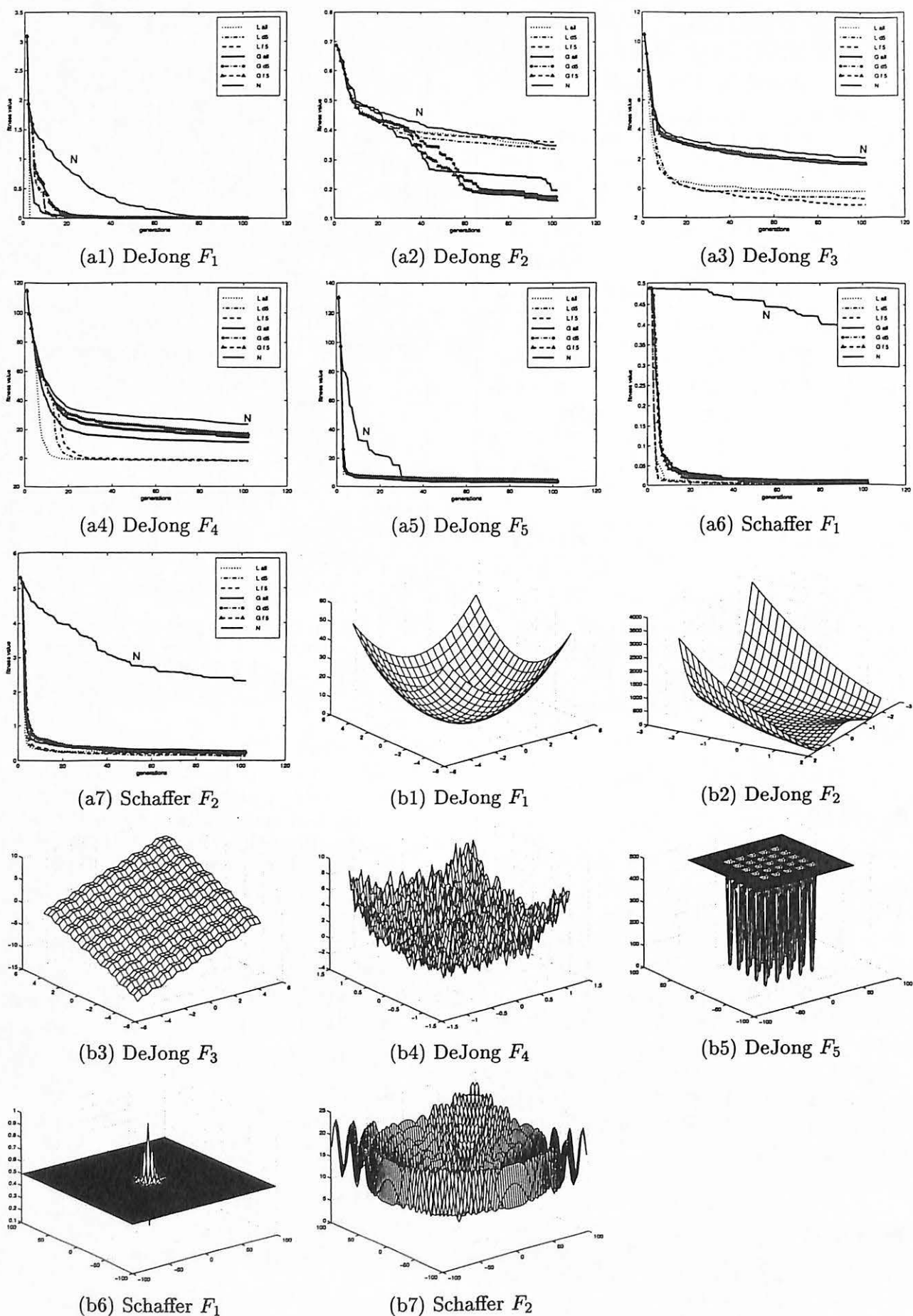


図 3: (a) GA 収束曲線と, (b) 各ベンチマーク関数の形状。記号 'N' を提案手法の収束曲線と区別するため、従来の GA の収束曲線に付した。図中の記号は表 1 を参照。

表 1: 性能比較実験に用いた手法。左の表の記号にある数字は、たとえば5の時、 $5 * (2n + 1)$ のデータが二次形式の関数生成に使用されていることを意味する。 n は次元数である。

L_{all}	提案手法。LS法で、すべてのデータを使用して二次形式の関数を生成。
L_{f5}	提案手法。LS法で、評価値基準データ選択法で二次形式の関数を生成。
L_{d5}	提案手法。LS法で、エリート距離基準データ選択法で二次形式の関数を生成。
G_{all}	提案手法。SubGA法で、すべてのデータを使用して二次形式の関数を生成。
G_{f5}	提案手法。SubGA法で、評価値基準データ選択法で二次形式の関数を生成。
G_{d5}	提案手法。SubGA法で、エリート距離基準データ選択法で二次形式の関数を生成。
N	提案手法を用いていない従来のGA。

いる。単峰性関数によるあてはめの計算時間を短縮するために初期の世代にのみ本手法を適用したり、数世代毎に適用するなどすれば、計算コストと性能の比がさらに良くなることが考えられる。

5 結論

単峰性曲面を探索空間にあてはめ、得られた曲面の極小値を新しいエリート個体として用いる手法を提案した。実験結果から、ベンチマーク7関数すべてに対して有意に効果を示すことができた。

本手法は得られたエリート個体と現世代の個体1つだけを入れ替える。したがって、得られたエリート個体がよくなかったとしても、我々の手法を組み込んでない従来のGAと性能は変わらない。我々の実験結果にはこのような特徴が現れている。これは、本提案手法のもう一つの利点といえよう。

参考文献

- [1] K.A. De Jong. An analysis of the behavior of a calaa of genetic adaptive systems. *Doctoral Dissertation, Univ. of Michigan, Univ. Microfilms*, Vol. 76, No. 9381, 1975.

表 2: GA の設定。

メイン GA	
コーディング方法	実数コーディング
世代数	100
個体数	30
選択方法	ルーレットホイール選択 エリート戦略
交叉方法	2点交叉
交叉率	90 %
突然変異方法	ガウシアン
突然変異率	5% per chromosome
SubGA	
コーディング方法	バイナリコーディング
個体数	20
世代数	5
選択方法	ルーレットホイール選択 エリート戦略
交叉方法	2点交叉
交叉率	90 %
突然変異方法	一様乱数
突然変異率	5% per bit

- [2] J.D. Schaffer, R.A. Caruana, L.J. Eshelmanm, and R. Das. A study of control parameters affecting online performance of genetic algorithms for function optimization. *3rd Int'l. Conf. on Genetic Algorithms (ICGA '89)*, 1989. Morgan Kaufmann Publishers, San Mateo, California.
- [3] 印具毅雄, 高木英行, 大崎美穂. 対話型遺伝的アルゴリズムのインターフェイス改善. 第13回ファジィシステムシンポジウム, pp. 859-862, Jun 1997. 富山.
- [4] 高木英行, 畠見達夫, 寺野隆. 対話型進化計算法の研究動向. 人工知能学会誌, Vol. 13, No. 5, pp. 692-703, 1998.

問い合わせ先

〒815-8540

福岡市南区塩原4丁目9-1

九州芸術工科大学

高木英行

Tel&Fax: 092-553-4555

E-mail: takagi@kyushu-id.ac.jp