2009-ICS-154 (24)
2009/3/4

実数値GAにおけるスキーマ保存を考慮した交叉方法の提案

早瀬 智英† 松田 聖‡

†日本大学大学院生産工学研究科〒275-8575 千葉県習志野市泉町 1-2-1

‡日本大学生産工学部〒275-8575 千葉県習志野市泉町 1-2-1

E-mail: †sacredsugar@cnc.jp, ‡matsuda.satoshi@nihon-u.ac.jp

あらまし 実数値 GA に対する新しい交叉方法として、スキーマ保存を考慮した交叉方法を提案する。一般に、GA の有効性の拠り所は良質のスキーマが世代に渡って保存・進化されることである。しかし、実数値 GA に対してこれまで提案されている多くの手法はスキーマ保存をあまり配慮しておらず、GA の持つ有効性を活かしきいれていないのではないかと考える。そこで、スキーマ保存を考慮した新しい交叉方法を提案し、いくつかの数値実験を通してその有効性を検証する。

キーワード 実数値 GA, スキーマ, ビルディングブロック仮説

Schema Oriented Crossover for Real-Coded GA

Tomohide HAYASE† and Satoshi MATSUDA‡

† Department of Mathematical Information Engineering, Graduate School of Industrial Technology,
‡ Department of Mathematical Information Engineering, College of Industrial Technology,

Nihon University, 1-2-1 Izumi, Narashino-shi, Chiba, 275-8575 Japan
E-mail: †sacredsugar@cnc.jp, ‡matsuda.satoshi@nihon-u.ac.jp

Abstract A new crossover method for real-coded GA is proposed, which intends to preserve and evolve schema. We think that the major principle of GA as a search algorithm is to emerge, preserve, and evolve schema, which has a good part of solution. However, the methods for real-coded GA proposed so far seem weak in this aspect. The performance comparisons are made through simulations with other existing methods.

Keyword real –coded genetic algorithms, schema, building block hypothesis

1. はじめに

遺伝的アルゴリズムはナップザック問題に代表される組合せ最適化問題などの２値問題において高い有効性を示している。しかし、実際にはビットストリングで表現できる問題は少なく、パラメータが連続的な値をとる実数値の問題にも対応する必要がある。その中で実数を遺伝子表現に用いる実数値遺伝的アルゴリズムは非線形関数の大域的最適化手法として期待されており、そのための交叉法が多く提案されている[1][2][3]。

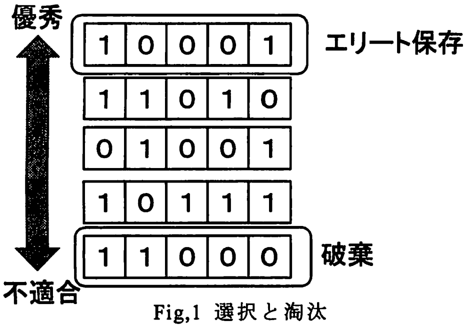
単峰性正規分布交叉などのやり方が高い効果を持つことはすでに報告されているが、これらの方法はスキーマの保存があまり考慮されておらず、GA の有効性を生かしきれていないと考えられる。一般に、GA の有効性の拠り所は良質のスキーマが世代に渡って保存・進化されることである。

そこで本論文ではスキーマ保存を考慮した交叉方法を提案し、いくつかの数値実験からその有用性を示す。２章では遺伝的アルゴリズムについてその概要を説明し、3 章では本論で提案する交叉方法を説明する。４章ではテスト関数を用いた数値実験の結果を示す。最後に５章は本稿のまとめを述べる。

2. GA 遺伝的アルゴリズム

遺伝的アルゴリズム(GA)とは、近似解を探索するためのメタヒューリスティックアルゴリズムである。遺伝的アルゴリズムは解の候補 x を遺伝子として表現した「個体」を複数用意し、適応度の高い個体を優先的に選択して、交叉・突然変異などの操作を繰り返しながら解を探索する。適応度は適応度関数によって与えられる。適応度関数によって得られた適応度に対し降順にソートを行う。その結果、優秀な適応度を持つ個

体はエリート遺伝子であるとして、次世代へ保存される。反対に適応度の低い個体は不適合であるとして破棄される。それ以外の個体は、後述する交叉や突然変異の操作を行い、次世代の遺伝子を作るための親遺伝子となる。



Fig,1 選択と淘汰

交叉とは、個体の遺伝子の一部を入れ替える操作である。遺伝子列から一点、または複数の点を選び、その交叉点において選択した二つの親遺伝子の配列内容を交換する操作である。この操作により、前世代遺伝子の特徴が引き継がれた次世代遺伝子が生成される。下記に一般的なアルゴリズムを示す。

① 初期世代遺伝子をn個生成する
② 評価関数を用いて生成した個体の適応度を計算する
③ 個体を二つ選択し、交叉を行い新たな遺伝子をn個生成する。
④ 一定確率で突然変異を行う
⑤ ③で保持した遺伝子と親世代を降順にソートし、優秀方から遺伝子n個を子世代とする
⑥ 最も優秀な遺伝子をエリートとして保存する
⑦ ②～⑥の操作を最大世代まで繰り返し、最終的に最も適応度の高い個体を解とする。

Fig.2 は一点交叉の例を示す。

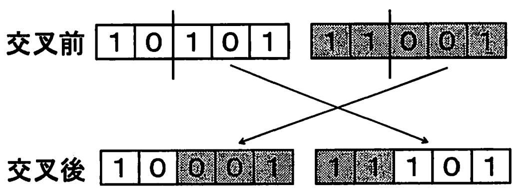


Fig.2 一点交叉

交叉とエリート保存を繰り返していくと、類似した遺伝子ばかりになってしまうことがある。その状況を避けるために突然変異という操作が用いられる。この操作は、交叉の操作とは別に一定確率で遺伝子の一部をランダムに書き換える。この操作により、遺伝子の多様性を保つことができる。Fig.3 に突然変異の一例を示す。

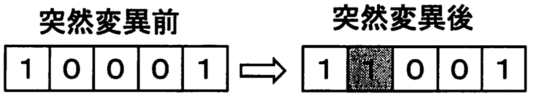


Fig.3 突然変異

上記の操作を繰り返し行い、徐々にある環境に適応した特徴を次世代に遺伝していき、世代を重ねるごとに環境に適した遺伝子を残すことが可能である。

3. スキーマ保存を考慮した交叉

本研究で提案する手法はスキーマ保存を考慮したものである。通常の交叉を用いた遺伝的アルゴリズムを行い、個体のもつ遺伝子の形質を残した上で多様性を持たせようとする提案である。

冒頭で述べたように、一般にGAの有効性の拠り所は良質のスキーマが世代に渡って保存・進化されていることである。しかし、実数値GAに対してこれまで提案されている多くの手法はスキーマ保存に配慮していない。UNDX[4]等の単峰性正規分布交叉法[小野 99]・REX等の大域的降下方向を用いた提案[小林 03]、などはビットストリングGAと異なり、遺伝子間の値を交換するのではなく距離・位置関係から一定の空間をとり、その中から子を生成する手法をとっている。Fig.4に単峰性正規分布交叉法の概念図を示す。

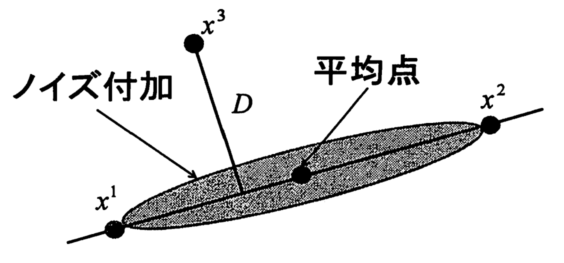


Fig.4 単峰性正規分布交叉概念図

これに対し、スキーマ保存を考慮した交叉では通常の交叉を行い、生成した子の遺伝子が持つ値に一定のランダム値を与えて変化を促すというものである。一点交叉では、親の良質のスキーマは交叉点で分断されない限り子に継承される、また、子の値をランダムに変化させることにより、親遺伝子の持つスキーマを残しつつ、その周辺を探索することを可能としている。Fig.5 にスキーマ保存を考慮した一点交叉の概念図を

示す。

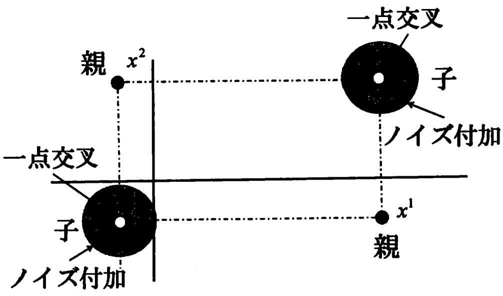


Fig.5 概念図

このランダムに変化する値は、世代の進行に伴い値を小さくする。これは適応解の値が小さくなるため、変化する値が大きいと解が収束しないためである。
ノイズ値の変化を定めている式を以下に示す。

$$10^{-n}x \quad n = 0,1,2,\ldots$$

x はノイズの初期値であり、今回は 0.1 で実験を行った。n の値は関数が最適な値に近づくにつれて値を増やすことでノイズの変化量を抑制する。今回の実験では 10^{-n} の値を遺伝子の適合度が下回ったとき、n の値を 1 増加させた。この n の増加速度を調節することでノイズ値を抑制する速度を調節することができる。

なお、交叉点で良質のスキーマが分断され、子に継承されない場合でも、交叉後に親遺伝子を残すことにより、集団の中から消滅しないようにしている。

4. 性能評価シミュレーション

本提案手法の最適値探索能力を評価するために実数値 GA の性能評価によく用いられている６つのテスト関数[1]に適用し、既存の手法と比較を試みた。

4.1 テスト関数

下記にそれらのテスト関数について説明する。

・Sphere 関数

原点を最小点とする二次関数。全ての x が０のとき最小値０をとる。変数間の依存関係は無い。

$$f_{Sphere}(x) = \sum_{i=1}^{n} x_i^2$$

$$-5.12 \le x_i \le 5.12, \quad 1 \le i \le n$$

・Rastrigin 関数

二次関数に三角関数が重畳された多峰性の強い関数であるが、変数間の依存関係のない分離型の関数である。全ての x が０のとき最小値０をとる。

$$f_{Rastrigin}(x) = 200 + \sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi x_i)\right]$$

$$-5.12 \le x_i \le 5.12, \quad 1 \le i \le n$$

・Ellispsoid 関数

目的関数に対する感度が変数によって大きく異なる性質である、悪スケール問題を持つ関数。比較的弱い悪スケールを持つ。全ての x が０のとき最小値０をとる。

$$f(x) = \sum_{i=1}^{n}\left(1000^{i-1/n-1} x_i\right)^2$$

$$-5.12 < x_i < 5.12, \quad 1 \le i \le n$$

・k-tablet 関数(k=n/4)

比較的強い悪スケール問題を持つ問題。全ての x が０のとき最小値０をとる。

$$f(x) = \sum_{i=1}^{k} x_i^2 + \sum_{i=k+1}^{n}(100 x_i)^2$$

$$-5.12 < x_i < 5.12, \quad 1 \le i \le n$$

・Bohachevsky 関数

比較的弱い多峰性を持つ関数。全ての x が０のとき最小値０をとる。

$$f(x) = \sum_{i=1}^{n-1}\ (x_i^2 + 2x_{i+1}^2 - 0.3\cos(3\pi x_i) - 0.4\cos(4\pi x_{i+1}) + 0.7)$$

$$-5.12 \le x_i \le 5.12, \quad 1 \le i \le n$$

・Rosenbrock 関数

単峰性であるが深く曲がった谷間の中に最適解があり、変数間の相互依存関係が強い非分解な関数である。(1,...,1)で最小値０をとる。

$$f_{Rosebrock}(x) = \sum_{i=2}^{n}\left[100\left(x_1 - (x_i)^2\right)^2 + (x_i - 1)^2\right]$$

$$-2.048 \le x_i \le 2.048, \quad 1 \le i \le n$$

4.2 シミュレーション結果と評価

これらの関数に対し、スキーマ保存を考慮した交叉方法と、文献[1]に示された、単峰性正規分布交叉(UNDX)・REXの性能を比較する。

次元は3つの関数全て20次元(n=20)で、集団サイズはRastrigin関数は親世代150個・子世代300個で、それ以外の関数は全て親世代25個・子世代50個で実験を行った。

また、本研究では最良個体の評価値が10^{-7}以下になったとき最適解到達と判定し、30試行すべてが最適解に収束するとき安定収束するとした。

以下にスキーマ保存を考慮した交叉方法と、UNDX[2]・e-UNDX[1]そして・REX-star[1]との比較を以下のTableに示す。Tableには収束するまでに要した平均評価回数と、比較手法の本提案手法に対する百分率を表している。

(1)UNDXとの比較

関数	本提案手法(a)	UNDX(b)	a/b(%)
Sphere	1.51×10^4	1.72×10^5	8%
Rosenbrock-star	1.02×10^5	1.20×10^6	9%
Rastrigin	3.79×10^4	3.41×10^6	1%未満

Table.1　UNDXとの比較(収束までの評価回数)

UNDX[2]との比較においては、3つの関数全てにおいて10%以下の速度で収束し、良好な性能を示している。

(2)スキーマ保存を考慮した交叉方法とe-UNDX

関数	本提案手法(a)	e-UNDX(b)	a/b(%)
Sphere	1.51×10^4	3.03×10^4	50%
Ellisp	収束せず	4.20×10^4	-
k-tablet	収束せず	7.10×10^4	-
Rosenbrock-star	1.02×10^5	2.10×10^5	49%
Boha	5.82×10^4	3.77×10^4	154%
Rastrigin	3.79×10^4	3.07×10^5	12%

Table.2　e-UNDXとの比較(収束までの評価回数)

e-UNDX[1]との比較においては、Spher関数では50%・Rosenbrok関数では約49%、Rastrigin関数では15%の速度で収束し、良好な性能を示しているといえる。しかしEllipsoid関数・k-tablet関数においてはスキーマを考慮した交叉方法では収束せず、悪スケール問題に対して非常に弱いという性質を持っていると推測できる。

(3)REX-starとの比較

関数	本提案手法(a)	REX-star(b)	a/b(%)
Sphere	1.51×10^4	6.89×10^3	219%
Ellisp	収束せず	8.46×10^3	-
k-tablet	収束せず	1.05×10^4	-
Rosenbroc k-star	1.02×10^5	5.45×10^4	187%
Boha	5.82×10^5	1.54×10^4	3779%
Rastrigin	3.79×10^4	1.23×10^5	31%

Table.3 REX-starとの比較(収束までの評価回数)

REX-star[1]との比較においては、Rastrigin関数のみREX-starを上回っている。他の収束している関数においても、性能が下回っている。

これらのことから、スキーマを考慮した交叉方法は、多峰性のある関数・変数間に関係のある式に対しては優良な性能を発揮する。中でもRastrigin関数に対しては非常に高い性能を持つ。しかし悪スケールを持つ関数に対しては現状の段階では対抗できず、改善が必要である。

また、遺伝子の値を変化させるノイズ値、ノイズ値の減少速度は収束速度に大きな影響を与える。今のところ、関数ごとに有効な値・減少速度は異なり、手動で調節しなくてはならない。

なお、比較した既存の手法は長年にわたって改良がなされて、提案当初から大幅に性能が向上しているが、[5], [6], [7]本提案手法は現段階では基本アイディアを単純に実装したものであり、具体的な実装方法には多くのバリエーションがあり、ちょっとした変更で性能がかなり左右されることがある。今後はこのような実装方法に対する改良、又、現実的な問題への適応を通じての評価を行っていきたい。

5. おわりに

本研究では、実数値GAにおけるスキーマ保存を考慮した交叉を提案し、その有用性について、いくつかのテスト関数への適用を通じて、その性能を検証した。ノイズ値の決定、悪スケール問題への対処などの問題を抱えているが、多峰性を持つ関数・変数間の関係が強い関数に対しては良好な性能を示すことを紹介した。

今後の課題として、悪スケール問題への対処を可能とする手法の考案、現在手動で行っているランダムに変化するノイズ値の設定を自動で決定できるようにすることによる汎用的な交叉法の確立、などが考えられ

る。これらの問題を解決し、性能の向上を試みる。

文　献

[1] 小林　重信, "実数値 GA のフロンティア", 人工知能学会誌, 24 巻, 1 号, 128p～143p.　January 2009.

[2] 小野　功 "実数値 GA のための正規分布交叉の多数の親を用いた拡張法の提案"　計測自動制御学会論文 36 巻, 10 号, pp.875-883 2000.

[3] 小野　功 "実数値 GA とその応用"　人工知能学会誌 ,15 巻 ,2 号, pp.259-266, March 2000.

[4] 小野　功, "単峰性正規分布を用いた実数値 GA による関数最適化" 人工知能学会誌, 14 巻, 6 号, 1146p～1155p.　1997

[5] 木村　周平, "交叉の設計指針に基づく UNDX の拡張 : ENDX の提案と評価" 計測制御学会論文集 36 巻, 12 号, pp.1162-1171, 2000

[6] 喜多　一, "実数値 GA のための正規分布交叉のお多数の親を用いた拡張法の提案", 計測制御学会論文集 36 巻, 10 号, pp.875-883, 20009ccCc

[7] 樋口　隆英 "実数値 GA におけるシンプレクス交叉の提案", 人工知能学会論文誌, 16 巻, 1 号, 2001