

系統的サンプリング手法を備えた遺伝的アルゴリズムの開発

Development of Genetic Algorithms with Systematic Sampling

柴田 圭一郎

Kei-ichiro SHIBATA

阿部 充志

Mitsushi ABE

(株)日立製作所 電力・電機開発研究所

Hitachi, Ltd. Power & Industrial Systems R & D Laboratory

An efficient technique, which improves diversity of population in Genetic Algorithms (GAs) for global optimizing problem, has been developed, based on multi-dimensional Low-Dispersion Sequences (LDS) samplings. The LDS enables uniformly distributed sampling in each dimension. This characteristics improves the diversity of the population and increases variable dimension up to 30, applying to Real-Coded GAs, which have fine convergence to the optimum solution, on complicated benchmark function (Schwefel's function).

1. はじめに

遺伝的アルゴリズム(Genetic Algorithms; GA)は生物の進化に着想を得た最適化手法であり, 解集合を保持しながら探索する確率的多点探索法である. 伝統的なビットストリング GA における収束性能を改善すべく, 近年, 実数値ベクトルを用いてコード化した実数値 GA が提案され, 様々な交叉手法が検討されている[小野 00]. 広域での探索性能と収束性能は相反する要求であり, この課題に対して, 集団の多様性を維持するため, 低分散列[Niederreiter 92]を用いて探索領域から系統的に個体をサンプリングする手法を備えた実数値 GA を提案する.

2. 系統的サンプリング手法を備えた実数値 GA

2.1 低分散列を用いた系統的サンプリング

(1) 低分散列の構成方法

乱数よりも一様かつ稠密で, 区間(0,1)に生成される点列を準乱数と呼び, モンテカルロ法による高次元積分において高精度の積分値を得るために用いられている[Niederreiter 92]. 準乱数の1種である低分散列の構成方法を以下に示す. 基底 $b \geq 2$ 及びサンプル番号 $n \geq 1$ を任意の整数とし, m を $\log_b(n)$ の整数部分とすると, n は b のべき展開係数 a_i を用いて次式で表される.

$$n = \sum_{i=0}^m a_i b^i = a_0 + a_1 b + \dots + a_m b^m \quad \dots (1)$$

n : サンプル番号(任意の正整数)

b : 基底(2以上の任意の正整数)

a_i : 基底 b のべき展開係数

次に, 生成行列と呼ばれる $m \times m$ の正方行列に, 式(1)における基底 b のべき展開係数ベクトル $\mathbf{a} = (a_0, a_1, \dots, a_m)^T$ を乗ずることにより k 次元におけるべき展開係数ベクトル $\mathbf{a}(k) = (a_0(k), a_1(k), \dots, a_m(k))^T$ を求める. 実用上生成可能な次元数が40次元程度と手頃なことから, 今回の探索では, Pascal行列[手塚 03]と呼ばれる上三角行列から算出される生成行列を用いるScrambled Faure列(以下SF列と記す)を採用した. 次元数以上の最小素数基底 b の剰余系を用いて式(2)により k 次元のべき展開係数ベク

トル $\mathbf{a}(k)$ を求め, 得られた結果を式(3)に代入することにより, k 次元のSF列 $\Phi_b^{(k)}(n)$ を得ることができる.

$$\mathbf{a}(k) = (\mathbf{P}^T)^{k-1} \mathbf{P}^{k-1} \mathbf{a} \quad \dots (2)$$

\mathbf{P} : Pascal行列

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 0 & 1 & 2 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

$$\Phi_b^{(k)}(n) = \sum_{i=0}^m \frac{a_i(k)}{b^{i+1}} = \frac{a_0(k)}{b} + \frac{a_1(k)}{b^2} + \dots + \frac{a_m(k)}{b^{m+1}} \quad \dots (3)$$

(2) 低分散列の特長

第一次元を横軸, 第二次元を縦軸にとり, 三次元SF列分布を投影した結果を図1に示す. 一見して729($=9^3$)個のサンプル点が一様に分布しており, 詳細に見るとこのサンプル点は区間(0,1)を9等分した81個の部分区間に9点ずつ属している. SF列のこの規則性は横軸, 縦軸に任意の次元を選んだ場合にも成立している. 即ち, 729個のサンプル点は各次元を9等分した部分区間に1個ずつ属している. このような性質を持つ低分散列を実数値GAに適用することにより, 探索領域から系統的に漏れなく個体をサンプリングすることが可能となる.

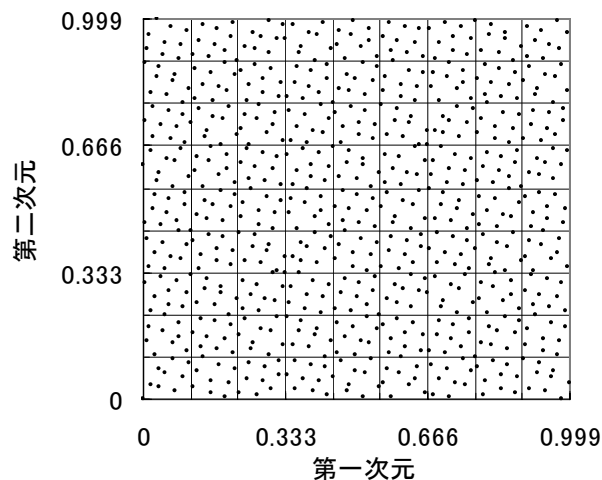


図1 三次元Scrambled Faure列

連絡先: 柴田圭一郎, (株)日立製作所 電力・電機開発研究所,
反応・計測システムプロジェクト,
〒319-1221 茨城県日立市大みか町 7-2-1,
0294-52-9316, keiichiro.shibata.pk@hitachi.com

2.2 実数値 GA への適用

低分散列を適用して探索領域を系統的にサンプリングする機能を備えた実数値 GA のフローチャートを図2に示す。初期個体集団から次元数+1個の親個体をランダムに選択してシンプレクス交叉 (Simplex Crossover; SPX) [樋口 01]により子個体を生成する。次に、2.1 節に示した SF 列を探索領域に線形変換して新個体を生成する。最小世代ギャップモデル (Minimal Generation Gap; MGG) [佐藤 97]において親個体及び子個体に、一定の比率で新個体を追加して世代交代を行う。新個体の適合度が高ければ、その個体は GA に組み込まれ、周囲が重点的に探索される。以上に示した方法により、広域での探索性能と収束性能を両立させることが可能となる。

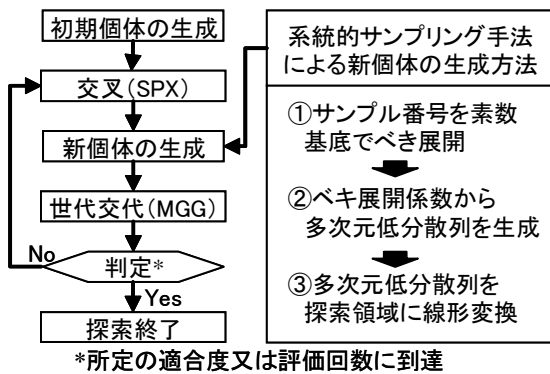


図2 系統的サンプリング手法を備えた実数値GAのフローチャート

3. ベンチマーク関数による性能評価

3.1 ベンチマーク関数及び計算条件

従来の SPX と MGG による実数値 GA (従来法) 及び系統的サンプリング手法を備えた実数値 GA (新個体法) について、表1に示すベンチマーク関数に適用してその性能を比較した。

各ベンチマーク関数の次元数 n を 10, 20, 30 とし、単谷性関数 (Scaled-Rosenbrock 関数) の初期個体数を次元数 $n \times 15$ 個、多谷性関数 (Rastrigin-1.0 関数, Schwefel 関数) の初期個体数を次元数 $n \times 90$ 個に設定した。又、SPX により生成する子個体数を次元数 $n \times 10$ 個に設定した。解の多様性を維持しつつ、解の収束を妨げないという条件でサーベイした結果から、新個体数を子個体数の 5% に設定した。GA の探索結果は乱数に依存し、常に最適解が得られるとは限らないので、関数値 $< 10^{-7}$ を収束条件として、乱数を変更して夫々 10 回探索した。

表1 ベンチマーク関数

関数名	関数式 探索領域
Scaled-Rosenbrock	$f(x) = \sum_{i=2}^n [100 \{x_i - (ix_i)^2\}^2 + (ix_i - 1)^2]$ $-2.048 \leq x_i \leq 2.048 \quad / i$
Rastrigin-1.0	$f(x) = 10n + \sum_{i=1}^n [(x_i - 1)^2 - 10 \cos \{2\pi(x_i - 1)\}]$ $-5.12 \leq x_i \leq 5.12$
Schwefel	$f(x) = An + \sum_{i=1}^n \{x_i \sin(\sqrt{ x_i })\}$ $-512 \leq x_i \leq 512, A = 418.9828872724$

3.2 探索結果

Scaled-Rosenbrock 関数及び Rastrigin-1.0 関数では、新個体法の収束に要する評価回数が、従来法の夫々 1.1 倍及び 1.7 倍程度に増加するものの、両手法とも最適解に収束した。

Schwefel 関数は、最適解が探索領域の端部にあり、最小の局所最適解と離れているため探索が困難である。10, 20, 30 次元の Schwefel 関数の探索結果を表2に示す。10 次元では、従来法は3割の成功率であるのに対して、新個体法は評価回数が 1.2 倍となるものの 10 割の成功率である。20 及び 30 次元では、従来法は最適解に到達できないが、新個体法は夫々 6 割及び 4 割の確率で最適解に到達した。

30 次元 Schwefel 関数において最小値を得た収束曲線を図3に示す。従来法では約 3×10^7 回で局所最適解に収束するのに対し、新個体法では、約 8×10^7 回までは緩やかに減少し、それ以後は急速に減少して約 10^8 回で 10^{-7} 以下の最適解を得た。

表2 Schwefel関数の探索結果

次元数	従来法 (SPX+MGG)		新個体法	
	成功率*1	評価回数*2	成功率*1	評価回数*2
10	3/10	2.3×10^6	10/10	2.7×10^6
20	0/10	-	6/10	2.1×10^7
30	0/10	-	4/10	1.1×10^8

*1 成功率: 最適解に到達した回数/試行回数

*2 評価回数: 最適解に到達した場合の平均値

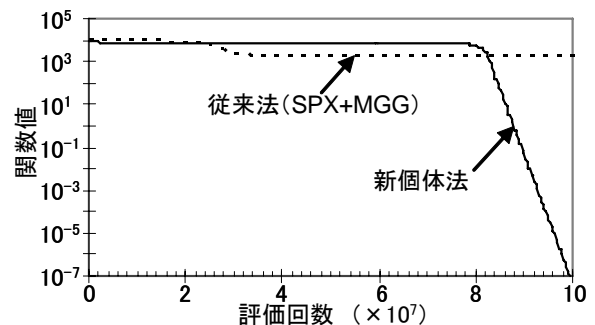


図3 30次元Schwefel関数の収束曲線

4. まとめ

- 低分散列を実数値 GA に適用して広域での探索性能と最適解への収束性能を両立させた系統的サンプリング手法を備えた実数値 GA を提案し、その性能を評価した。
- 複雑なベンチマーク関数 (Schwefel 関数) に対しても系統的サンプリング手法を備えた実数値 GA を用いることにより、30 次元以上の探索を可能とした。

参考文献

- [小野 00] 小野功, 山村雅幸, 喜多一, 実数値 GA とその応用, 人工知能学会誌 (2000)
- [Niederreiter 92] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAMS, Philadelphia (1992)
- [手塚 03] 手塚集, 他 4 名, 統計科学のフロンティア 11 計算統計 I - 確率計算の新しい手法 -, 岩波書店 (2003)
- [樋口 01] 樋口隆英, 筒井茂義, 山村雅幸, 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 (2001)
- [佐藤 97] 佐藤浩, 小野功, 小林重信, 遺伝的アルゴリズムにおける世代交代モデルの提案と評価, 人工知能学会誌 (1997)