

分散確率モデル遺伝的アルゴリズム

Distributed Probabilistic Model-Building Genetic Algorithm

非 佐野 正樹 (同志社大院) 正 廣安 知之 (同志社大工)
正 三木 光範 (同志社大工) 非 下坂 久司 (同志社大院)
非 筒井 茂義 (阪南大経営情報)

Masaki SANO, Graduate School of Engineering, Doshisha University
Tomoyuki HIROYASU, Doshisha University, Tatara Miyakodani 1-3, Kyo-Tanabe, Kyoto
Mitsunori MIKI, Doshisha University
Hisashi SHIMOSAKA, Graduate School of Engineering, Doshisha University
Shigeyoshi TSUTSUI, Hannan University

Algorithms where offsprings (new search points) are generated according to the estimated probability model of the good parents are called the Probabilistic Model-Building Genetic Algorithms (PMBGAs). In this paper, a new model of PMBGA, Distributed PMBGA (DPMBGA), is proposed. In the DPMBGA, the correlation between the design variables is considered by PCA when the offsprings are generated. The distribution of the offsprings is estimated as the normal distribution. The island model is also applied in the DPMBGA for maintaining the population diversity. Through the standard test functions, the effectiveness of the DPMBGA is examined. The result shows the good search ability of the DPMBGA with PCA for the test functions that have correlation between the design variables. On the other hand, the DPMBGA without PCA is good at optimizing the problems where there is no correlation between the design variables. The DPMBGA where PCA is executed in the half of the islands and not executed in the other island can find the good solutions in the problems whether or not the problems have the correlation between the design variables. The results of the DPMBGA are also compared with those of the UNDX with MGG. The results explain that the DPMBGA shows the better performance than the UNDX.

Key word: genetic algorithms, probabilistic model-building genetic algorithms, distributed genetic algorithm, principal component analysis

1 はじめに

遺伝的アルゴリズム (Genetic Algorithm : GA) は, 生物の進化と自然淘汰を工学的に模倣した最適化アルゴリズムである¹⁾. 遺伝的アルゴリズムでは, 探索空間上の探索点を生物の個体とみなす. 個体の母集団 (population) に対して, 選択 (selection), 交叉 (crossover), 突然変異 (mutation), という遺伝的操作 (genetic operator) を繰り返し適用する. 遺伝的操作の中でも, 各個体の持つ探索点の情報を交換する交叉は, GA の解探索における中心的な役割を担うものと考えられてきた. 交叉により, 良好な個体の持つ優れた部分解が結合し, より適合度の高い個体が生成されることが期待される. しかし, 同時に部分解を破壊する可能性もあることが知られている. この問題を解決するための新しいアプローチとして, 母集団内の良好な個体の統計情報を用いて新しい個体を生成する, 確率モデル遺伝的アルゴリズム (Probabilistic Model-Building Genetic Algorithm : PMBGA) についての研究が多くなされている²⁾.

本研究では, 新しい PMBGA のモデルである, 分

散確率モデル遺伝的アルゴリズム (Distributed PMBGA : DPMBGA) を提案する. DPMBGA では, 主成分分析 (Principal Component Analysis : PCA) によって親個体の分布を変換することにより, 設計変数間の依存関係を考慮して子個体を生成する³⁾. また, 個体の母集団を複数のサブ母集団に分割することで, 多様性の維持を図っている.

本論文の構成は, 以下のとおりである. 2 節と 3 において, 提案モデルのベースとなっている PMBGA と分散遺伝的アルゴリズムについて説明する. 4 節で, 提案モデルである DPMBGA について説明する. 5 節において, 数値実験を通じて提案モデルの特性および有効性について検討する.

2 確率モデル遺伝的アルゴリズム

PMBGA では, GA と同様にランダムに生成された個体群の中から, 良好な解が選択される. 選ばれた個体の確率分布が推定され, 確率モデルが構築される. 構築した確率モデルに従い, 新しい探索点が生成される. こうして生成された新しい探索点は母集

団内の個体と置き換えられる．これを終了条件を満たすまで繰り返す．したがって，PMBGA は，GA における交叉による子個体の生成を，次の2つのステップと置き換えたものであると考えることができる．

1. 選択された個体のモデル（分布の推定による）を構築する
2. 新しい探索点を構築したモデルに従って生成する

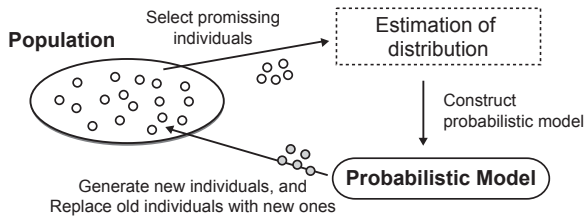


Fig. 1: Probabilistic Model-Building GA

3 分散遺伝的アルゴリズム

GA は，多点探索であり，評価計算を反復して行うため，計算コストが高いという問題点がある．このため，GA の並列モデルについては，多くの研究がなされてきた⁴⁾．

GA の並列モデルの1つに，Tanese によって提案された，分散遺伝的アルゴリズム（Distributed GA : DGA）がある⁵⁾．このモデルでは，個体の母集団を複数のサブ母集団（島）に分割し，島ごとに遺伝的操作を適用する．DGA は島モデルとも呼ばれる．また，一定世代ごとに他の島と個体を交換する．この個体の交換を移住 (migration) といい，移住を行う間隔と移住を行う個体数の割合を，それぞれ移住間隔，移住率という．DGA は通常の単一母集団モデルと比較して，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⁾．

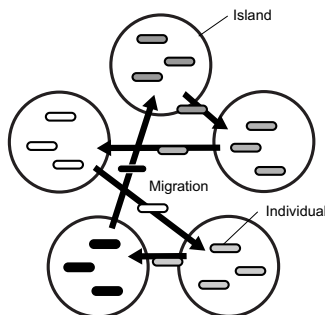


Fig. 2: Distributed Genetic Algorithm

4 分散確率モデル遺伝的アルゴリズム

本研究では，新しい PMBGA のモデルである分散確率モデル遺伝的アルゴリズム（Distributed Probabilistic Model-Building Genetic Algorithm : DPMBGA）を提案する．DPMBGA では，DGA と同様に母集団を複数の島に分割し，島ごとに PMBGA を実行する．また，確率分布のモデル構築の際に，主成分分析（PCA）によって個体群の分布を変換する．これにより，設計変数間の依存関係を考慮して子個体を生成することができる．

4.1 DPMBGA の概要

DPMBGA では，全個体を複数の島に分割する．各島内において PMBGA を行い，一定世代ごとに，島間で個体の交換（移住）を行う．本論文で採用するトポロジは，移住のたびに，全ての島が無作為な順番の1つのリングを形成し，隣の島が移住先となるものである．移住個体の決定法は，島内からランダムに選出された個体を送り出して最悪個体と入れ替える方法である．

DPMBGA では，次の手順を各世代 t において実行する（Fig. 3）．

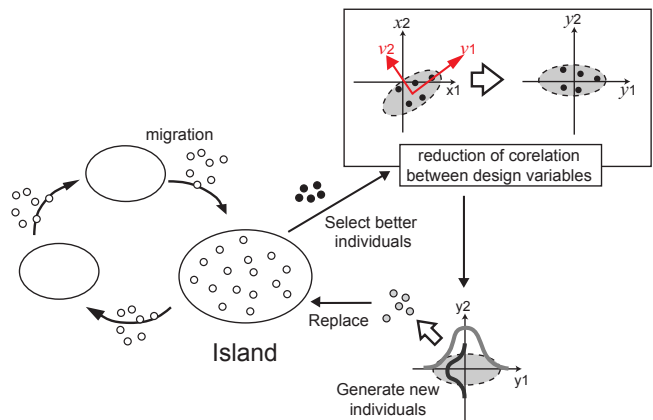


Fig. 3: DPMBGA

1. エリートの保存
2. 良好な個体の抽出
3. PCA を用いて，良好な個体群の設計変数無相関化
4. 新しい個体の生成
5. 設計変数の相関の復元と島内の個体の置き換え
6. 突然変異
7. 制約条件外の個体を，実行可能領域の境界上に引き戻し

8. エリートの復帰

9. 個体の適合度を評価

以降では、それぞれの操作について説明する。

4.2 良好な個体の抽出

各島 $P_{sub}(t)$ から、良好な個体を、抽出率 R_s で定められた割合だけ選択し、サンプル個体 $S(t)$ として抽出する。新しい子個体はこれらの情報を基に生成される。適合度の高い順番に個体を選択するが、同じ個体は重複して選択しない。重複個体を除いたときに島内に個体が不足する場合には、ランダムに個体を生成して $S(t)$ に追加する。.. $S(t)$ は各島ごとに存在する。

4.3 PCA を用いた、良好な個体群の設計変数の無相関化

$S(t)$ は PCA によりデータ変換される。PCA に利用するサンプル個体は $S(t)$ とは別に用意する。現在の世代までに島内に出現した最良個体群を、PCA の対象とする個体群 $T(t)$ とする (Fig. 4)。現世代までに島内に出現した個体の数が $T(t)$ のサイズに満たない場合、新しい個体を追加せずそのまま処理を続ける。 $T(t)$ のサイズを超えた場合には、劣個体から順に削除する。島内の個体数にかかわらず、任意の数の個体を PCA に用いることができる。これらのアーカイブは島ごとに存在するものとする。

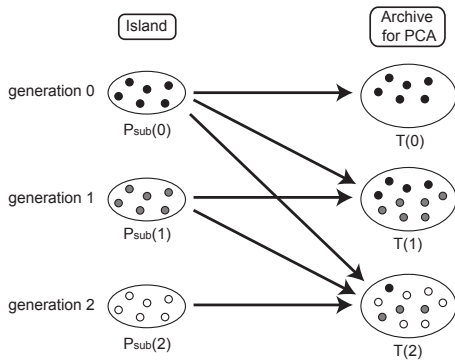


Fig. 4: PCA with the archive of the best individuals

$T(t)$ の各設計変数 (D 個) から $T(t)$ の平均値を引いたものを、行列 T ($nT(t)$ 行 $\times D$ 列) とする。 T は、 $T(t)$ の平均偏差行列である。抽出した $S(t)$ の各設計変数 (D 個) から $T(t)$ の平均値を引いたものを、行列 X ($nS(t)$ 行 $\times D$ 列) とする。

次に、 T の共分散行列 S (D 行 $\times D$ 列) を求め、その固有値と固有ベクトルを算出する。 S は、実数

値対称行列となり、次式で求まる。

$$S = \frac{1}{nS(t) - 1} T^T T \quad (1)$$

S の固有値は大きい順に $\lambda_1, \lambda_2, \dots, \lambda_D$ とし、対応する固有ベクトルを v_1, v_2, \dots, v_D (D 次元) とする。最も大きい固有値は、設計変数の分散が最大となるように座標軸をとったときの分散値に等しい。固有ベクトルは、その座標軸自身を示す。

求めた固有ベクトルを用いて、抽出した個体群 $S(t)$ の設計変数 X を無相関化する。固有ベクトルを並べて、座標軸変換のための行列 $V = [v_1, v_2, \dots, v_D]$ を作成する。 X に対して座標変行列 V をかけたものを、 Y とおく。 Y は、 $nS(t)$ 行 $\times D$ 列の行列であり、無相関化された設計変数を示す。 Y の座標軸は、固有ベクトルに等しい。

$$Y = XV \quad (2)$$

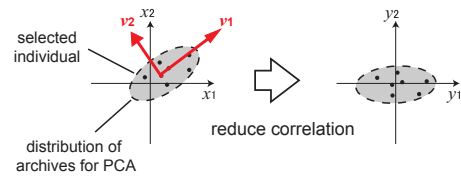


Fig. 5: Reduction operation of correlation between design variables with PCA

4.4 新しい個体群の生成

新しい子個体を島内の個体数と同じ数 ($nP(t)$) だけ発生させる。 Y の分布に従い、正規乱数によって、各設計変数を独立に決定する。すなわち、設計変数が n の場合には、 n 個の正規分布を用意する。正規乱数の分散は、 Y における各設計変数 (列) の分散に倍率 Amp をかけたものである。また、その平均は、 Y における各設計変数の平均に等しい。発生させた個体の設計変数は、 Y_{offs} ($nP(t)$ 行 $\times D$ 列) に格納する。

4.5 設計変数の相関の復元と島内の個体の置き換え

Y_{offs} に V の逆行列をかけ、座標軸を元に戻す。

$$X_{offs} = Y_{offs} \cdot V^{-1} \quad (3)$$

X_{offs} の平均値を元に戻し、島内の個体 $P(t)$ と入れ替えて $P(t+1)$ とする。

4.6 突然変異

あらかじめ定めた突然変異率 R_{mu} にしたがって、設計変数を、制約条件内の無作為な値に変更する。

4.7 エリートの保存と復帰

あらかじめ定めた数 ($nE(t)$) のエリートを $E(t)$ として保存する．確率モデルによる個体の生成の後，エリートを母集団に戻す $E(t)$ を，島内の個体群 $P(t+1)$ の劣悪な個体と置き換える．

4.8 特徴

DPMBGA の特徴は以下のとおりである．

- 実数値確率モデル GA である．
- 島モデルにより，多様性の維持を図る．
- 主成分分析 (PCA) を用い，1 次設計変数間の依存関係を考慮して子個体を生成する．
- 確率分布に正規分布を採用する．

DPMBGA では，良好な個体の分布を基に，正規分布によって新しい個体を発生させる．このため，現在の世代の良好な個体付近に収束する傾向が強く，早期収束が発生しやすいモデルであるといえる．そこで，局所解へ陥ることを防ぐため，母集団を複数の島に分割する島モデルを採用し，多様性の維持を図っている．

5 数値実験

5.1 対象問題

本論文で対象とするテスト関数は，以下に示す Rastrigin 関数，Schwefel 関数，Rosenbrock 関数，Ridge 関数，Griewank 関数の 5 つである．いずれも最小化問題であり，大域的最適値は 0 である．

Rastrigin 関数と Schwefel 関数は，設計変数間に依存関係の無い多峰性の関数である．Rosenbrock 関数と Ridge 関数は，設計変数間に依存関係のある単峰性の関数である．Griewank 関数は，設計変数間に依存関係のある多峰性の関数である．

$$F_{Rastrigin} = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)) \quad (4)$$

$(-5.12 \leq x_i < 5.12)$

$$F_{Schwefel} = \sum_{i=1}^n -x_i \sin(\sqrt{|x_i|}) - C \quad (5)$$

$(C : \text{optimum.})$
 $(-512 \leq x_i < 512)$

$$F_{Rosenbrock} = \sum_{i=2}^n (100(x_1 - x_i^2)^2 + (1 - x_i)^2) \quad (6)$$

$(-2.048 \leq x_i < 2.048)$

$$F_{Ridge} = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2 \quad (7)$$

$(-64 \leq x_i < 64)$

$$F_{Griewank} = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \left(\cos\left(\frac{x_i}{\sqrt{i}}\right) \right) \quad (8)$$

$(-512 \leq x_i < 512)$

5.2 主成分分析の効果と環境分散スキームの検討

DPMBGA では，主成分分析 (PCA) を用いて個体群の分布を変換することで，設計変数間の依存関係を考慮して新しい個体を生成する．設計変数間に依存関係を有する問題においては，個体の各設計変数の分布は，他の設計変数の影響を受ける．PCA を用いて設計変数の分布を無相関化することにより，親個体の分布が持つ性質を保存して子個体を生成することができる．よって，このような問題に対して，DPMBGA が良好な性能を示すことが期待される．

本節では，DPMBGA における PCA の効果について，数値実験を通じて検討する．数値実験で用いるパラメータは，Table 1 のとおりである．対象問題については，Schwefel 関数は 10 次元のものを，それ以外の関数は 20 次元のものを用いた．

Table 1: Parameters

Population size	512
Number of elites	1
Number of islands	32
Migration rate	0.0625
Migration interval	5
Archive size for PCA	100
Sampling rate	0.25
Amp. of Variance	2
Mutation rate	0.1/ (Dim. of function)

数値実験では，次に示すモデルについて比較を行う．

model 1 : 全ての島において PCA を行う．

model 2 : 全ての島において PCA を行わない．

model 3 : 全体の半分の島において PCA を行う．それ以外の島では行わない．

各モデルは，PCA を行う島の数異なる．model 1 は，4 節で述べたアルゴリズムと同様のものである．model 2 は PCA を全く行わないモデルである．

model 3 は PCA の使用に関して、環境分散スキームを採用している。

環境分散スキームを用いた GA (Distributed Environment GA : DEGA) は、三木らによって提案された DGA のモデルである⁶⁾。DEGA では、各島が異なるパラメータで遺伝的操作を行う。GA の解探索能力はパラメータ設定に依存する。最適なパラメータは対象問題に依存し、未知であるため、予備実験や専門的知識に基づいて経験的に設定しなければならない。これは、利便性の観点から好ましくない。そこで、DGA において各島に異なるパラメータを設定し、どのような問題に対しても安定した解探索性能を実現することを目的としたモデルが DEGA である。

関数評価値の閾値を $1 \cdot 10^{-10}$ 、探索の終了条件を関数評価 $3 \cdot 10^6$ 回とし、20 試行中で閾値に到達した回数を Table 2 に示す。高い割合で閾値に到達しているモデルほど、安定した解探索を行っているといえる。また、閾値に到達した試行における関数評価回数の平均を、Fig. 6 に示す。より少ない関数評価回数で閾値に到達するモデルほど、性能が高いといえる。

Table 2: Number of times that the threshold is reached

	model 1	model 2	model 3
Rastrigin	0	20	20
Schwefel	20	20	20
Rosenbrock	20	0	20
Ridge	20	20	20
Griewank	19	17	20

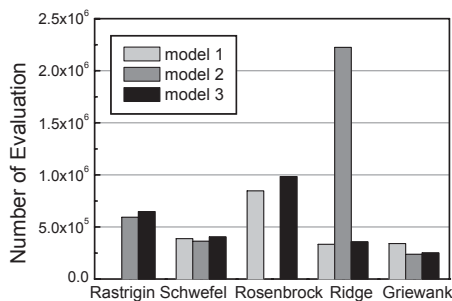


Fig. 6: Average number of evaluations to reach the threshold

Schwefel 関数では、どのモデルも、全ての試行において少ない評価回数で閾値に到達している。Rastrigin 関数では、model 1 の性能が悪い。Rastrigin

関数は多峰性の関数であるため、局所解に陥ることなく多様性を維持して探索を行うことが重要となる。PCA を行う model 1 では、設計変数間の相関を維持する方向にのみ個体を発生させるため、多様性が十分に維持されなかったと推測される。設計変数間に依存関係のある Rosenbrock 関数では、PCA を行わない model 2 が閾値に到達していない。同じく依存関係のある Ridge 関数では、model 2 は高い割合で閾値に到達しているが、多くの関数評価を必要としている。よって、これら関数に対しては、PCA が効率よく機能していると考えられる。以上の結果より、PCA による設計変数の無相関化は、設計変数間に依存関係のある単峰性の問題には有効に機能するが、依存関係の無い多峰性の問題に対しては十分な性能を実現できない場合があるといえる。

一方、PCA を行う島と行わない島との両方を有する model 3 は、どの関数に対しても優れた性能を示している。Griewank 関数は、設計変数間に依存関係があり、多数の局所解を持つ多峰性の関数である。このため、局所解に陥る可能性が高く、model 1 と model 2 との両方において、閾値に到達しない試行が存在する。これに対し、model 3 は、全ての試行において閾値に到達している。この結果より、model 3 は、設計変数の依存関係の有無に関わらず良好な解を発見し、多数の局所解を持つ問題に対しても安定した性能を示しているといえる。

5.3 UNDX を用いた MGG モデルとの比較

本節では、DPMBGA と既存の実数値 GA との比較を行う。比較対象は、単峰性正規分布交叉を用いた MGG モデルである。

単峰性正規分布交叉 (Unimodal Normal Distribution Crossover : UNDX) は、小野らによって考案された、実数値 GA の代表的な交叉法である⁷⁾。UNDX では、3つの親個体から2つの子個体が生成される。子個体の設計変数は、第1、第2の親個体が形成する主軸付近に、正規分布に従って決定される。第3の親は、主軸以外の成分の決定に用いられる。UNDX では、対象問題における設計変数間の依存関係に沿って子個体を生成することができる。

また、Minimal Generation Gap (MGG) は佐藤らによって考案された世代交代モデルである⁸⁾。MGG では、世代交代の際、母集団から親個体をランダムに選択する。親個体を1回あるいは複数回交叉させ、子個体を生成する。生成された子個体と親個体を合わせた個体群から、あらかじめ定めた選択法に基づいて次世代に生き残る個体を選択し、母集団に戻す。MGG では、世代交代の限定化と選択の

局所化とにより、多様性を維持して解探索を行うことが可能である。

Fig. 7 に、DPMBGA と比較対象のモデル (UNDX+MGG) との解の履歴を示す。横軸は目的関数の評価回数である。縦軸は関数評価値の 20 試行平均である。ただし、UNDX+MGG については、20 試行中で前述の閾値に到達した試行のみの平均値をプロットしている。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になる。

UNDX+MGG のパラメータについては、多峰性関数では 300 個体、単峰性関数では 50 個体とし、交叉回数 100 回、 $\alpha = 0.5$ 、 $\beta = 0.35$ 、としている。DPMBGA のパラメータ設定は、5.2 節の model 3 と同様である。

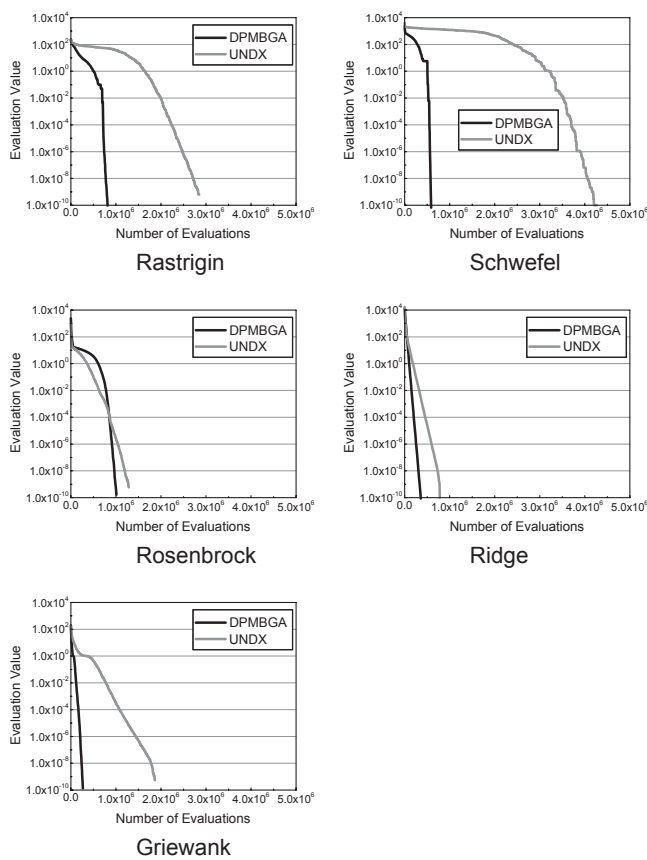


Fig. 7: History of average of the evaluation value

Fig. 7 より、どの関数に対しても、環境分散スキームを採用した DPMBGA が良好な性能を示している。よって DPMBGA は、設計変数間の依存関係の有無に関わらず、連続関数最適化において有効なモデルであるといえる。

6 まとめ

本研究では、確率モデル GA の新しいモデルである、分散確率モデル遺伝的アルゴリズム

(DPMBGA) を提案した。DPMBGA では、個体の確率分布のモデル構築に、主成分分析 (PCA) による設計変数の無相関化を導入することにより、設計変数間の依存関係を考慮して子個体を生成することができる。また、島モデルの採用により、多様性の維持を図っている。

テスト関数に対する数値実験の結果、次のことが明らかとなった。PCA を行う DPMBGA は、設計変数間に依存関係のある問題に有効である。PCA を行わない DPMGA は、依存関係の無い問題に有効である。PCA を行う島と行わない島とを含んだ環境分散モデルは、設計変数間の依存関係の有無に関わらず、安定して良好な解を得ることができる。また、既存の実数値 GA である、UNDX を用いた MGG モデルとの性能比較を行った。比較実験においては、本論文で使用した全てのテスト関数に対し、DPMBGA が良好な性能を示した。よって、DPMBGA は連続関数最適化において有効なモデルであるといえる。

謝辞

本研究は文科省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行った。ここに謝意を表する。

参考文献

- 1) D.E.Goldberg. *Genetic Algorithms in Search Optimization and Machine Learning*. Addison-Wesley, 1989.
- 2) Martin Pelikan, David E. Goldberg, and Fernando Lobo. A survey of optimization by building and using probabilistic models. No. 99018, Sep. 1999.
- 3) 高橋仁人, 喜多一. 実数値遺伝的アルゴリズムのための独立成分分析を用いた交叉の提案. 第 13 回 自律分散システム・シンポジウム 資料, pp. 245–250, 2001.
- 4) Erick Cantú-Paz. A survey of parallel genetic algorithms. *Calculateurs Paralleles*, Vol. 10, No. 2, 1998.
- 5) Reiko Tanese. Distributed genetic algorithms. *Proc. 3rd International Conference on Genetic Algorithms*, pp. 434–439, 1989.
- 6) M.Miki, T.Hiroyasu, M.Kaneko, and K.Hatanaka. A parallel genetic algorithm with distributed environment scheme. *IEEE Proceedings of Systems, Man and Cybernetics Conference SMC'99*, 1999.
- 7) 小野功, 佐藤浩, 小林重信. 単峰性正規分布交叉 UNDX を用いた実数値 GA による関数最適化. 人工知能学会誌, Vol. 14, No. 6, pp. 1146–1155, 1999.
- 8) 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, Vol. 12, No. 5, pp. 734–744, 1997.

出典：

日本機械学会 第5回最適化シンポジウム講演論文集，
No.02-03，
pp.65-70，
(2002年10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>)