

# 分散確率モデル遺伝的アルゴリズム

廣安知之<sup>†</sup> 三木光範<sup>†</sup> 下坂久司<sup>††</sup>  
佐野正樹<sup>††,☆</sup> 筒井茂義<sup>†††</sup>

本論文では、新しい確率モデル GA (PMBGA) である、分散確率モデル遺伝的アルゴリズム (DPMBGA) を提案する。DPMBGA では、主成分分析 (PCA) により、設計変数の相関を考慮して子個体を生成する。また、島モデルの採用により、多様性の維持を図っている。テスト関数を用いた数値実験により、DPMBGA の有効性を検証した。複数のモデルを検討した結果、全体の半数の島でのみ PCA を行う DPMBGA は、対象問題における設計変数間の依存関係の有無にかかわらず、良好な性能を示した。解探索性能は、単峰性正規分布交叉を用いた Minimal Generation Gap モデルと比較して DPMBGA がより優れた性能を示した。また、DPMBGA に Boundary Extension by Mirroring (BEM) を適用したモデルを用いて、探索領域の境界の取り扱いについても検討している。

## Distributed Probabilistic Model-Building Genetic Algorithm

TOMOYUKI HIROYASU<sup>†</sup>, MITSUNORI MIKI<sup>†</sup>  
, HISASHI SHIMOSAKA<sup>††</sup>, MASAKI SANNO<sup>††,☆</sup>  
and SHIGEYOSHI TSUTSUI<sup>†††</sup>

In this paper, a new model of Probabilistic Model-Building Genetic Algorithms (PMBGAs), Distributed PMBGA (DPMBGA), is proposed. In the DPMBGA, the correlation among the design variables is considered by Principal Component Analysis (PCA) when the offsprings are generated. 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. In this paper, some models of DPMBGA are examined. The DPMBGA where PCA is executed in the half of the islands and not executed in the other islands can find the good solutions in the problems whether or not the problems have the correlation among the design variables. From these results, it is clarified that the DPMBGA has higher searching ability than the Unimodal Normal Distribution Crossover with Minimal Generation Gap. It is also discussed the treatment of the boundary condition of the design field using the Boundary Extension by Mirroring (BEM).

### 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<sup>1)</sup>。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。これらの操作の中で、効率良く探索を行うためには、交叉において親の持つ良質な形質が子が受け継ぐことが一つの方法である。これを行う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手法が挙げられる。これらの手法は、確率モデル遺伝的アルゴリズム (Probabilistic Model-Building Genetic Algorithm : PMBGA)<sup>2)</sup> もしくは Estimation of Distribution Algorithm (EDA)<sup>3)</sup> と呼ばれ、多くの研究がなされている。

本研究では、連続関数最適化に注目する。連続関数の最適化においては、目的関数において、設計変数間に依存関係がある場合、良好な解を得るためには、依存関係を考慮して探索点を生成することが良好な探索につながる。設計変数値をそのまま個体の染色体として用いる実数値 GA (Real-Coded GA) では、こ

<sup>†</sup> 同志社大学工学部

Department of Knowledge Engineering and Computer Sciences, Doshisha University

<sup>††</sup> 同志社大学大学院

Graduate School of Engineering, Doshisha University

<sup>☆</sup> 現在, IBM Japan, Ltd.

<sup>†††</sup> 阪南大学経営情報学部

Department of Management Information, Hannan University

の点に留意した交叉法の研究がなされている。代表的なものに、小野らによって考案された単峰性正規分布交叉 (Unimodal Normal Distribution Crossover : UNDX)<sup>4)</sup>がある。UNDXでは、子個体の設計変数が2つ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。UNDXは、設計変数間に依存関係を有する問題に対して良好な解を得ることができる。しかし、UNDXによって発生した子個体は多様性を欠く。これに対し、高橋らにより、独立成分分析を用いた実数値交叉<sup>5)</sup>が提案されている。この手法では、主成分分析と独立成分分析で座標系を変換した後に、Eshelmanらによるブレンド交叉 (Blend Crossover : BLX- $\alpha$ )を適用する。多様性の維持に優れたBLX- $\alpha$ を、個体分布を変換してから適用することで、依存関係の考慮と多様な解の生成との両方を図っている。また、GAと類似した進化計算手法に、進化戦略<sup>6)</sup>がある。Schwefelによって提案された correlated mutationでは、個体の分布の方向を示すパラメータが採用されており、設計変数間の依存関係を考慮した突然変異を行う。

一方、連続関数最適化において設計変数間の依存を考慮したPMBGAは少ない。そこで、本論文では、連続関数最適化に対する新しい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。先に述べた、進化戦略における correlated mutation<sup>6)</sup>では、設計変数間の相関の情報を各個体が保持する。そして、突然変異や組み替えによって、確率的にその適切な値を探索する。これに対し、DPMBGAでは、良好な個体を蓄積したアーカイブにPCAを適用し、相関の情報を算出する。また、DPMBGAでは、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。DPMBGAの利点として、PCA変換を行う集団を別に採用することで、より設計変数間の依存情報を個体生成の際に反映することが可能となること、個体生成の際に確率モデルを利用することで、より親の形質を子に伝えることが可能となることなどが挙げられる。

本論文の構成は、以下のとおりである。2節と3節において、提案モデルのベースとなっているPMBGAと分散確率モデル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

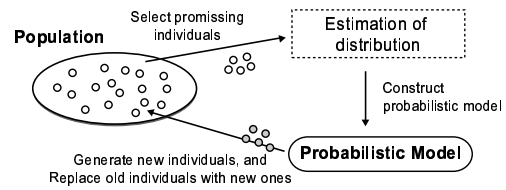


図1 確率モデル遺伝的アルゴリズム  
Fig. 1 Probabilistic Model-Building GA

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。図1はこの流れを模式的に表したものである。したがって、PMBGAは、GAにおける交叉による子個体の生成を、良好な個体の確率モデルの構築と、構築したモデルによる新しい探索点の生成とに置き換えたものであると考えることができる。

## 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<sup>7)</sup>。

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<sup>8)</sup>。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用を適用する。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<sup>8)</sup>。本研究では、DGAを解の多様性を維持するメカニズムとして利用している。

## 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

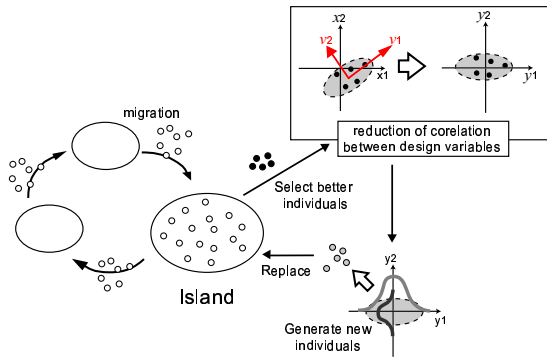


図2 分散確率モデル遺伝的アルゴリズム  
Fig. 2 DPMBGA

に母集団を複数の島に分割し、島ごとに PMBGA を実行する。母集団の分割により、多様性が維持され、複雑な問題においても局所解に陥ることなく探索を行うことが期待される。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、母集団内の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 (図2)。

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を無相関化
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。 $S(t)$  は島ごとに存在する。新しい子個体は

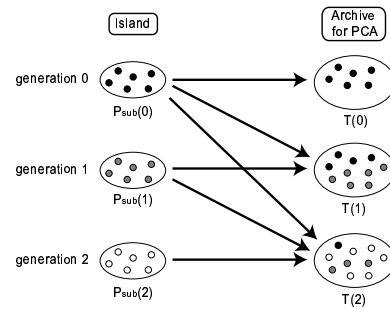


図3 PCA への過去の最良個体群の利用  
Fig. 3 PCA with the archive of the best individuals

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、同じ設計変数値を持つ個体は重複して選択しない。このことは、抽出される個体数が少ないことに起因する。DPMBGA では多くのサブ母集団を用いて探索を行うため、一島あたりの母集団サイズは必然的に小さくなる。そのため、重複個体を選択することは、確率分布を推定するための情報量が大きく減少することを意味する。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。追加する個体の設計変数値は、定義域内の一様乱数によって設定する。

#### 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 (図3)。以降、この  $T(t)$  をアーカイブと呼ぶ。アーカイブの更新は毎世代行われる。アーカイブには常に適合度の高い順に個体が格納されるため、各世代に更新される個体数は探索の状況によって異なる。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にそのまま処理を続行する。 $T(t)$  のサイズを超えた場合には、適合度の低い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。

アーカイブに格納されている個体数を  $N_T$  とし、 $T(t)$  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  $\mathbf{T}$  ( $N_T$  行  $\times D$  列) とする。サンプル個体数を  $N_S$  とし、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  $\mathbf{X}$  ( $N_S$  行  $\times D$  列) とする。

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

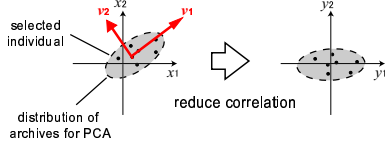


図4 座標変換による、設計変数値の無相関化

Fig. 4 Reduction operation of correlation between design variables with PCA

対称行列となり、次式で求まる。

$$\mathbf{S} = \frac{1}{N_T - 1} \mathbf{T}^T \mathbf{T} \quad (1)$$

$\mathbf{S}$  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  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_D$  ( $D$  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  $\mathbf{X}$  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  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_D]$  を作成する。 $\mathbf{X}$  に対して座標軸変換行列  $\mathbf{V}$  をかけた行列を、 $\mathbf{Y}$  とおく。 $\mathbf{Y}$  ( $N_S$  行  $\times$   $D$  列) は、無相関化された設計変数を示す。 $\mathbf{Y}$  の座標軸は、固有ベクトルに等しい。

#### 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 ( $N_P$ ) だけ発生させる。 $\mathbf{Y}$  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、 $\mathbf{Y}$  における各設計変数 (各列) の分散に倍率  $Amp$  をかけたものである。また、その平均は、 $\mathbf{Y}$  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、 $\mathbf{Y}_{offs}$  ( $N_P$  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

$\mathbf{Y}_{offs}$  に  $\mathbf{V}$  の逆行列をかけ、座標軸を元に戻す。

$$\mathbf{X}_{offs} = \mathbf{Y}_{offs} \cdot \mathbf{V}^{-1} \quad (2)$$

$\mathbf{X}_{offs}$  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 ( $N_E$ ) 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、多様性の維持を図っている。
- 主成分分析 (PCA) を用い、1 次の設計変数間の依存関係を考慮して子個体を生成する。
- 最良個体を蓄積したアーカイブを用いることで、島内の個体数にかかわらず、任意の数の個体を PCA の適用対象とすることができる。
- 確率分布に正規分布を採用している。

DPMBGA では、新しい個体は、良好な個体の分布を基に、正規分布によって生成される。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なり、早期収束が発生しやすいモデルであるといえる。そこで、母集団を複数の島に分割する島モデルを採用している。母集団の分割により、多様性が維持され、複雑な問題に対しても、局所解へ陥ることなく探索が行われることが期待される。

## 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。Schwefel 関数は 10 次元のものを、それ以外の関数は 20 次元のものをを用いる。

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。

$$F_{Rastrigin} = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)) \quad (3)$$

$$(-5.12 \leq x_i < 5.12)$$

$$F_{Schwefel} = \sum_{i=1}^n -x_i \sin(\sqrt{|x_i|}) - C \quad (4)$$

$$(C : optimum.)$$

$$(-512 \leq x_i < 512)$$

$$F_{Rosenbrock} = \sum_{i=2}^n (100(x_1 - x_i^2)^2 + (1 - x_i)^2) \quad (5)$$

$$(-2.048 \leq x_i < 2.048)$$

表 1 パラメータ  
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)

$$F_{Ridge} = \sum_{i=1}^n \left( \sum_{j=1}^i x_j \right)^2 \quad (6)$$

$(-64 \leq x_i < 64)$

$$F_{Griewank} = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \left( \cos\left(\frac{x_i}{\sqrt{i}}\right) \right) \quad (7)$$

$(-512 \leq x_i < 512)$

## 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では、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、表 1 のとおりである。

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<sup>9)</sup>。DEGA では、各島が異なるパラメータで遺伝的操作を行う。GA の解探索における最適なパラメータは対象問題に依存し、未知である。そこで、DGA において各島に異なるパラメー

表 2 20 試行中で最適値に到達した回数

Table 2 Number of times that the threshold is obtained

	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

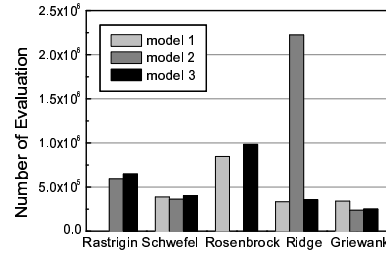


図 5 最適値に到達するまでの関数評価回数の平均

Fig. 5 Average number of evaluations to reach the threshold

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。

本検討では、関数評価値が  $1.0 \times 10^{-10}$  以下のとき、最適値に到達したとみなす。探索の終了条件を関数評価  $3.0 \times 10^6$  回とし、20 試行中で最適値に到達した回数を表 2 に示す。高い割合で最適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、最適値に到達した試行における関数評価回数の平均を、図 5 に示す。より少ない関数評価回数で最適値に到達するモデルほど、性能が高いといえる。

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最適値に到達している。Rastrigin 関数では、model 1 の性能が悪い。すなわち、PCA が探索性能を悪化させている。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 が最適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 は高い割合で最適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、最適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最適値に到達している。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 主成分分析が有効に機能しない原因についての検討

5.2 節の数値実験において、Rastrigin 関数に対しては、PCA の実行が探索性能を悪化させることが明らかとなった。PCA が有効に機能しない原因として、以下のことが考えられる。DPMBGA では、PCA が最良個体のアーカイブに対して適用されるため、望ましい探索性能を得るためには、アーカイブの個体分布が対象問題の特性を正確に反映したものである必要がある。しかし、局所解を多く持つ Rastrigin 関数においては、多くの個体が局所解に陥り、アーカイブが更新されにくくなる可能性がある。このような場合には、アーカイブに最新の探索結果が反映されなくなり、PCA によって適切な個体分布の変換が行われないと考えられる。

上記の点について、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、表 1 の通りである。検討対象とする DPMBGA のモデルは、5.2 節の model 1（全ての島で PCA を行うモデル）である。対象問題は、Rastrigin 関数と Rosenbrock 関数である。

図 6 に、アーカイブの更新の履歴を示す。横軸は目的関数の評価回数である。縦軸は更新された個体数の 20 試行平均である。この値が大きいほど、その世代の多くの個体がアーカイブに導入されたことを示す。同図より、Rastrigin 関数では、探索が進むとアーカイブの更新がきわめて少なくなる。一方、Rosenbrock 関数では、つねにアーカイブ内の多くの個体が更新されていることがわかる。

図 7 に、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。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。図中の normal は通常モデルであり、erase/10 は、10 世代ごとにアーカイブ内の個体を全て消去するモデルである。同図より、アーカイブを 10 世代ごとに消去すると、Rastrigin 関数では探索性能が向上し、逆に、Rosenbrock 関数では解探索の効率が悪化する。このことは、Rastrigin

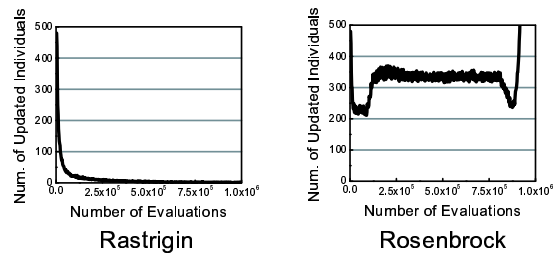


図 6 最良個体アーカイブにおける更新された個体数の履歴  
Fig. 6 History of number of updated individuals in archive of the best individuals

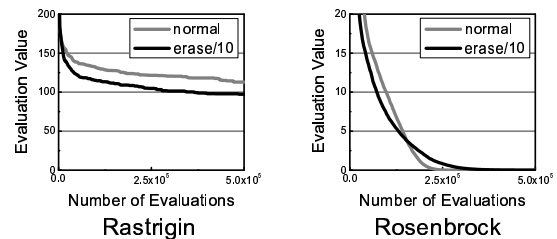


図 7 10 世代ごとにアーカイブを消去するモデルにおける解の履歴  
Fig. 7 History of average of evaluation values in the model in which archive is erased each 10 generation

関数において、アーカイブが有効に機能していないことを示している。

以上の結果より、Rastrigin 関数において PCA が探索性能を悪化させる原因の 1 つは、アーカイブの更新の停滞であるといえる。更新の停滞は、アーカイブ内の個体が局所解を含む峰に陥ることに起因すると推測される。先に述べたとおり、アーカイブを用いると、実際の個体数が少ない場合でも、多くの個体を PCA の対象とすることができるという利点がある。しかし、本節の実験により、設計変数間に依存関係の無い多峰性の関数においては、アーカイブの使用が探索性能を低下させることもあることが明らかとなった。

### 5.4 アーカイブサイズの検討

DPMBGA では、現在の世代までに島内に出現した最良個体群を PCA の対象とするアーカイブとして格納する。アーカイブに PCA を適用して得られる情報は、サンプル個体群の設計変数無相関化するための座標軸変換行列である。このため、アーカイブサイズは、PCA によって得られる座標軸変換行列の精度に大きく関係するパラメータであると考えられる。

本節の数値実験では、アーカイブサイズが解探索性能に与える影響を検討する。数値実験で用いたパラメータは表 1 と同様である。ただし、アーカイブサイズを 5, 20, 50, 100, 500 の 5 通りに変化させた。対象とする問題は、20 次元の Rastrigin 関数および

表 3 20 試行中で最適値に到達した回数

Table 3 Number of times that the threshold is reached

Size	Rastrigin		Rosenbrock	
	model 1	model 3	model 1	model 3
5	2	20	0	0
20	0	20	18	20
50	0	20	20	20
100	0	20	20	20
500	0	20	20	20

Rosenbrock 関数である。検討対象とする DPMBGA のモデルは、5.2 節の model1(全ての島で PCA を行うモデル)と model3(環境分散スキーム)である。探索の終了条件を関数評価  $3.0 \times 10^6$  回とし、20 試行中で最適値に到達した回数を示す。結果を表 3 に示す。

表 3 より、アーカイブサイズが 50 以上の場合、5.2 節の結果と同様に、Rastrigin 関数では model3 が、Rosenbrock 関数では model1 と model3 が安定した解探索性能を示している。しかしながら、アーカイブサイズが 50 より小さい場合、5.2 節の結果とは異なる傾向を示す。まず、Rosenbrock 関数においては、model1 と model3 の両方で最適解発見率が 0/20 となった。また、Rastrigin 関数の model1 において、アーカイブサイズが 5 の場合、最適解発見率は 2/20 となった。5.2 節の結果と照らしあわせると、アーカイブサイズが小さい場合、model2(全ての島で PCA を行わないモデル)と類似する傾向を示すことがわかった。また、20 次元の問題を対象とした場合、5.2 節の実験で用いたアーカイブサイズ 100 は、PCA が有効に機能するための十分なサイズといえる。

### 5.5 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<sup>4)</sup>。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<sup>10)</sup>。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

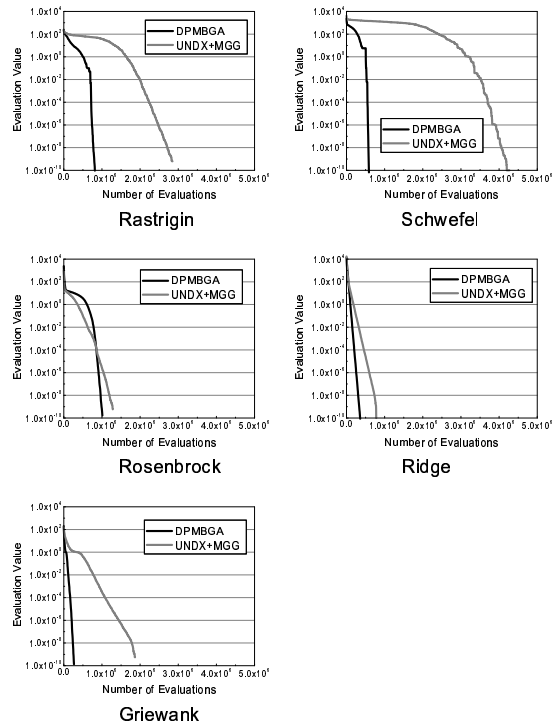


図 8 DPMBGA と UNDX を用いた MGG モデルとの比較  
Fig. 8 History of average of evaluation values

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。

図 8 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 試行中で前述の最適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、文献<sup>4)</sup>を参考に、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 (環境分散スキーム)と同様である。

図 8 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。

表 4 目的関数の定義域

Table 4 Domain of objective functions

Function	Optimum of $x_i$	Domain
Rastrigin	0.0	[0, 5.12]
Schwefel	420.968746	[-512, 421]
Rosenbrock	1.0	[-2.048, 1]
Ridge	0.0	[0, 64]

### 5.6 最適解が設計変数空間の端に位置する問題についての検討

実数値 GA の交叉法は、最適解が探索領域の中心に位置する問題に対しては良好な性能を示すが、探索領域の境界付近に最適解を有する問題に対しては良好な解を発見できないことが知られている。このような問題を解決する方法として、Tsutsui によって提案された Boundary Extension by Mirroring (BEM)<sup>11)</sup> がある。

BEM では、探索領域の境界から一定距離以内の個体の存在を許す。外部の個体を許容する範囲は、拡張率 (extension rate)  $r_e$  ( $0.0 < r_e < 1.0$ ) によって定義される。探索領域外の個体の関数評価値は、境界を基点として鏡像のように折り返した目的関数を用いて計算する。

DPMBGA は実数値 GA であるので、上記と同様の問題を有する可能性がある。また、探索領域の端では、個体の分布が探索領域の境界の作用を受けるため、確率モデル構築・子個体の生成のアルゴリズムの効果が、予期したものと異なることが推測される。本節では、探索領域の境界付近に最適解が位置する対象問題における DPMBGA の探索能力について検討する。BEM を DPMBGA に適用し、その有効性について議論する。DPMBGA のパラメータ設定は、5.2 節の model 3 (環境分散スキーム) と同様である。ただし、表 4 に示すように、最適解が探索領域の境界付近に位置するように定義域を設定している。

図 9 に、通常の DPMBGA および BEM を適用した DPMBGA 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。

同図より、BEM を用いない通常のモデルが、良好な性能を示している。DPMBGA では、探索領域の外に発生した個体については、最も近い探索領域の境界上に引き戻している。最適解が境界上にある場合、引き戻しによって境界付近に個体が集中する可能性がある。この点が、最適解が境界上にある関数において、BEM を適用しない通常のモデルがより良好な性能を示す原因であると考えられる。よって、探索領域が各

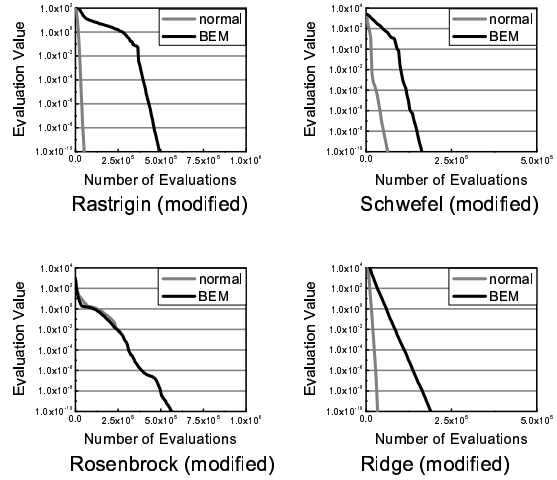


図 9 最適解が探索空間の端に位置する問題における解の履歴  
Fig. 9 History of average of evaluation values on functions which have optimum at the edge of search space

設計変数の上限と下限で定義されている問題については、最適解が端に位置していても、DPMBGA は良好な解を得ることができるといえる。

### 5.7 正規分布を用いた分布推定に関する検討

DPMBGA では正規分布により良好な個体群の分布を推定し、子個体を生成する。しかしながら、個体群の分布が一つの固まりではなく、複数の固まりから形成されるような場合には、個体群を正規分布で推定するには問題が生じる場合があるものと考えられる。例えば、多峰性の関数、または、最適解が端に位置するような関数において、これに該当する場合がある。

そこで、本節では 10 次元の Schwefel 関数を対象とし、上記の点について数値実験により分布を観察する。Schwefel 関数は多峰性の関数であり、かつ最適解が各設計変数の定義域の端付近に存在し、準最適解が他の端に存在する問題である。そのため、解の分布が複数に分離して存在する可能性がある。数値実験で用いる基本的なパラメータは、表 1 の通りであるが、分割母集団モデルを適用せず、PCA を行う 1 つの母集団 128 個体を対象とし、分散の増幅率を 1.0 とした。数値実験は、抽出された良好な個体群 (32 個体) と、生成された子個体群 (128 個体) の分布を比較することで行う。これらのパラメータは、良好な個体群と生成された個体群の分布の比較を行い易くするために決定した。分布の比較には、各世代における各設計変数の平均値と最適解までの距離の絶対値、分散、歪度を用いた。歪度とは分布の歪みを示す尺度であり、正方向への分布が伸びている場合は正、逆に負方向に伸びている場合は負の値をとる。また 0 に近ければ対称形に



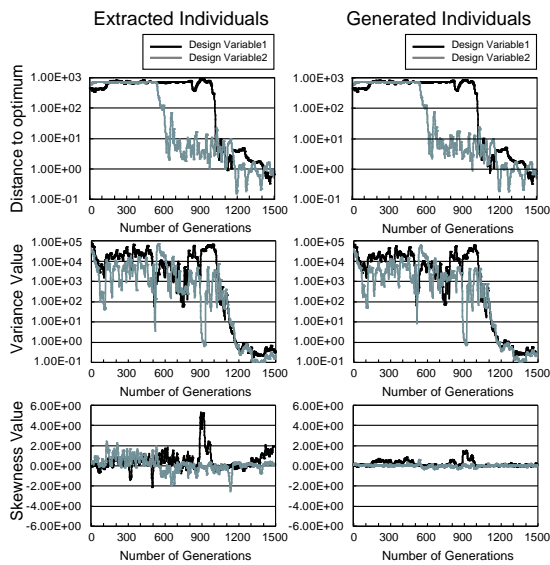


図 10 各設計変数の平均と最適解までの距離，分散，歪度の履歴  
 Fig. 10 History of distance between average and optimum point, variance value and skewness value of each design variable

近い分布であることを示している。図 10 に 1 試行の結果を示す。この図では、10 の設計変数のうち、2 つの設計変数の結果を示しており、10 世代で平滑化している。

図 10 より、1000 世代付近で設計変数 1 の最適解と個体群の平均の距離が急速に縮まっている。また、その直前で歪度が増加している。これらのことより、次のことが予想される。すなわち、1000 世代直前までは、局所解を探索していたのに対し、突然変異などにより、最適解付近の解が見つかり、そのため個体群の分布が大きくひずむ。その後、選択によって分布が最適解付近に遷移し、最適解が探索されているのである。これらの予想により、正規分布は推定モデルの候補の一つであるが、良好な個体群が複数の個体群に分割されているような場合には、正規分布による推定は適当ではないと考えられる。しかしながら、本研究の提案モデルは、分割母集団モデルであり、一つの分割母集団内の推定は、一つの正規分布によって行われているが、全体としては、複数の正規分布を利用して探索を行うモデルであるとも言える。そのため、良好な個体群が複数の個体群に分割されているような場合に対しても、単一母集団よりも推定しやすいものと考えられる。

## 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。

テスト関数を用いた数値実験の結果、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。アーカイブサイズの検討では、PCA を有効に機能させるためには、ある一定サイズ以上のアーカイブが必要なことがわかった。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。さらに、最適解が探索領域の端にある問題について検討を行った。このような問題に有効とされている BEM を DPMBGA に適用したモデルについて実験を行った結果、BEM を適用しない通常モデルが良好な性能を示した。また、個体生成の際に利用している正規分布の検討も行った。以上の結果より、DPMBGA は連続関数最適化において有効なモデルの一つであるといえる。

今後は DPMBGA の使用の際に必要ないくつかのパラメータの解に与える影響を検討する必要がある。

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) Goldberg, D.E.: *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-Wesley (1989).
- 2) Pelikan, M., Goldberg, D. E. and Lobo, F.: *A Survey of Optimization by Building and Using Probabilistic Models*, Technical Report 99018, IlliGAL (1999).
- 3) Larranaga, P. and Lozano, J.: *Estimation of Distribution Algorithms. A New Tool for Evolutionary Computation*, Kluwer Academic Publishers (2001).
- 4) 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).

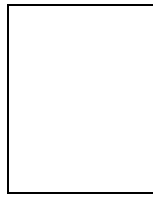
- 5) 高橋仁人, 喜多一: 実数値遺伝的アルゴリズムのための独立成分分析を用いた交叉の提案, 第13回自律分散システム・シンポジウム 資料, pp. 245-250 (2001).
- 6) Bäck, T., Hoffmeister, F. and Schwefel, H.-P.: A Survey of Evolution Strategies, *Proc. 4th International Conference on Genetic Algorithms*, pp. 2 - 9 (1991).
- 7) Cantú-Paz, E.: A survey of parallel genetic algorithms, *Calculateurs Paralleles*, Vol.10, No.2 (1998).
- 8) Tanese, R.: Distributed Genetic Algorithms, *Proc. 3rd International Conference on Genetic Algorithms*, pp. 434 - 439 (1989).
- 9) Miki, M., Hiroyasu, T., Kaneko, M. and Hatanaka, K.: A Parallel Genetic Algorithm with Distributed Environment Scheme, *IEEE Proceedings of Systems, Man and Cybernetics Conference SMC'99* (1999).
- 10) 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).
- 11) Tsutsui, S.: Multi-parent Recombination in Genetic Algorithms with Search Space Boundary Extension by Mirroring, *Proc. the 5th International Conference on Parallel Problem Solving from Nature (PPSN V)*, pp. 428-437 (1998).

(平成 xx 年 xx 月 xx 日受付)  
(平成 xx 年 xx 月 xx 日採録)



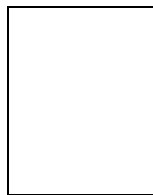
廣安 知之 (正会員)

1966 年生。1997 年早稲田大学理工学研究科後期博士課程終了。2001 年より同志社大学工学部専任講師。進化的計算, 最適設計, 並列処理, 設計工学などの研究に従事。IEEE, 電気情報通信学会, 情報処理学会, 日本機械学会, 計測自動制御学会, 超並列計算研究会, 日本計算工学会各会員。



三木 光範 (正会員)

同志社大学工学部知識工学科教授。現在の研究テーマは, 並列分散処理に基づくシステムの最適化, 遺伝的アルゴリズムやシミュレーテッドアーニーリングなどの進化的最適化手法の分散並列化, PC クラスタ・コンピューティングなど。1950 年生。1978 年大阪市立大学大学院工学研究科博士課程終了, 工学博士。大阪市立工業研究所研究員, 金沢工業大学助教授を経て 1987 年大阪府立大学工学部航空宇宙工学科助教授, 1994 年同志社大学工学部教授。進化的計算手法とその並列化, および知的なシステムの設計に関する研究に従事。著書は「工学問題を解決する適応化・知能化・最適化法」(技法堂出版) など多数。IEEE, 米国航空宇宙学会, 人工知能学会, システム制御情報学会, 日本機械学会, 計算工学会, 日本航空宇宙学会各会員。超並列計算研究会代表。



下坂 久司

1979 年生。2002 年同志社大学工学部卒業。同年同志社大学大学院工学研究科修士課程入学。Grid 計算環境における大規模最適化問題の進化計算手法による解法に興味を持つ。



佐野 正樹

1978 年生。2001 年同志社大学工学部卒業。2003 年同志社大学大学院工学研究科修士課程卒業。同年, 日本アイ・ビー・エム株式会社入社。ヒューリスティック最適化手法の 1 つである遺伝的アルゴリズムに興味を持つ。



筒井 茂義 (正会員)

1967 年大阪市立大学工学部電気工学科卒業。1969 年同大学大学院工学研究科修士課程修了。同年 (株) 日立製作所入社, 中央研究所, システム開発研究所勤務。現在, 阪南大学経営情報学部 経営情報学科, 兼大学院 企業情報研究科 教授。