

ヒストグラムを用いた確率モデル遺伝的アルゴリズムの提案

筒井茂義^{*1}, Martin Pelikan^{*2}, and David E. Goldberg^{*2}

*1 阪南大学経営情報学部, 〒580-8502 大阪府松原市天美東5-4-33, Phone: 072-332-1224, Fax: 072-336-2633, e-mail: tsutsui@hannan-u.ac.jp

*2 Illinois Genetic Algorithms Laboratory, Department of General Engineering, University of Illinois at Urbana-Champaign, 117 Transportation Building, 104 S. Mathews Avenue Urbana, Illinois 61801, USA, e-mail: {pelikan, deg}@illigal.ge.uiuc.edu

遺伝的アルゴリズム(以下, GA)における中心的なオペレータである「交叉」は, 短いサブ問題から構成される問題や規模が小さい問題においては, ビルディングブロックを効率よく組み合わせることが出来るのに対して, 複雑・大規模な問題では, 有益なビルディングブロックを破壊してしまうという問題がある. 近年このような従来のGAの問題を解決する方法として確率モデルGA(PMBGAs: probabilistic model-building genetic algorithms)の研究が活発に行われている. 確率モデルGAでは, 子個体は, 集団の確率分布に基づいて生成される. 本論文では, 確率モデルGAの一手法として周辺分布ヒストグラムを用いる方法を提案する. ヒストグラムモデルとしては, 等幅ヒストグラム(FWH: Fixed width histogram)と等高ヒストグラム(FHH: Fixed height histogram)を用いる. 実験の結果, 両モデルとも, パラメータ間に絡みがないかあるいは絡みが弱い問題では優れた探索特性をもっていることが明らかになった. 今後に残された課題についても考察する.

A Proposal of Probabilistic Model-Building Genetic Algorithms using Histogram Models

Shigeyoshi Tsutsui^{*1}, Martin Pelikan^{*2}, and David E. Goldberg^{*2}

*1 Department of Management Information, Hannan University, 5-4-33, Amaminihigashi, Matsubara, Osaka 580-8502 Japan

*2 Illinois Genetic Algorithms Laboratory, Department of General Engineering, University of Illinois at Urbana-Champaign, 117 Transportation Building, 104 S. Mathews Avenue Urbana, Illinois 61801, USA, e-mail: {pelikan, deg}@illigal.ge.uiuc.edu

Recently, there has been a growing interest in developing evolutionary algorithms based on probabilistic modeling. In this scheme, the offspring population is generated according to the estimated probability density model of the parents instead of using recombination and mutation operators. In this paper, we propose an evolutionary algorithm using a marginal histogram to model the parent population in a continuous domain. We propose two types of marginal histogram models: the fixed-width histogram (FWH) and the fixed-height histogram (FHH). The results showed that both models worked fairly well on test functions with no or weak interactions among variables. Especially, FHH could find the global optimum with very high accuracy effectively and showed good scale-up with the problem size.

1. はじめに

遺伝的アルゴリズム(以下, GA)における中心的なオペレータである「交叉」は, 短いサブ問題から構成される問題や規模が小さい問題においては, ビルディングブロックを効率よく組み合わせることが出来るのに対して, 複雑・大規模な問題では, 有益なビルディングブロックを破壊してしまうという問題がある.

近年このような従来のGAの問題を解決する方法として確率モデルGA(PMBGAs: probabilistic model-building genetic algorithms)の研究が注目をあびている. 確率モデルGAでは, 子個体は, 集団の確率分布に基づいて生成される. 確率モデルGAでは, 従来のGAと同様先ず最初ランダムに生成された個体から選択オペレータにより有望な個体を選択される. 次に, 選択された個体の確率分布の推定が行われ

る．この推定された確率分布に基づいて，次世代の子個体が生成（サンプリング）される．この操作が終了条件が満たされるまで繰り返される．従って，確率モデルGAは，交叉オペレータが確率分布の推定とそれに基づくサンプリングに置き換わっている点以外は，従来のGAと同じであり，各種の選択法や世代交代モデルを用いることが出来る．

確率モデルGAは，バイナリコーディングGAでの研究が先行している．変数（ビット）間の絡みを考慮する度合いにより，(1) no interactions, (2) pairwise interactions, (3) multivariate interactionsの三つに大きく分類することが出来る [Pelikan 00b]．変数間の絡みを考慮しない(1)の確率モデルGAでは，各変数は独立として扱われる．従って，このGAは，変数間に絡み合いがない問題に対しては，効率よい探索が期待できる．この例としては，PBIL(population based incremental learning)アルゴリズム[Baluja 94]，cGA(compact genetic algorithm) [Harik 98]，UMDA(univariate marginal distribution algorithm)[Mühlenbein 96]がある．(2)のpairwise interactionsでは，2変数間の絡みをモデルに取り込む．この分類の例としては，MIMIC(mutual-information-maximization input clustering)アルゴリズム[De Bonet 97]，2変数間の絡みをツリー状でモデル化するアルゴリズム[Baluja 97]がある．(3)のmultivariate interactionsでは，多変量間の絡みをモデル化する．このため，モデルの構築に大きな計算量が必要となるが，変数間に複雑な絡みを有する問題では，優れた探索性能を持っている．この分類の例としては，ECGA(extended compact genetic algorithm) [Harik 99]や BOA(Bayesian optimization algorithm) [Pelikan 99, 00a] がある．

実数値コーディングへの確率モデルGAの研究も近年行われている．PBILアルゴリズムを実数値に拡張する[Sebag 98]では，変数間の絡みが考慮されていない．変数間の絡みを考慮するEGNA(estimation of Gaussian networks algorithm)[Larranaga, 99]では，Gaussianネットワークにより，多変量正規分布のモデルが構築される．[Bosman 00b]では，クラスタリング手法と組み合わせて多変量混合正規分布を用い，非線形なパラメータ間の絡みを処理する方法を提案している．

本論文では，確率モデルGAの一手法として周辺分布ヒストグラムを用いる方法を提案し，関数最適化問題を用いた実験結果について述べる．ヒストグラムのモデルとしては，等幅ヒストグラム（FWH: Fixed width histogram）と等高ヒストグラム（FHH: Fixed height histogram）の二つを用いる[Tsutsui 01]．実験の結果，両モデルとも，パラメータ間に絡みがないあるいは絡みが弱い問題では優れた探索特性をもっていることが明らかになった．以下，先ず第2章において，FWHとFHHの二つのヒストグラムモデルについて述べる．次に，第3章において，両モデルを用いた確率モデルGAの実験結果について述べる．最後に第4章において，今後の課題について考察する．

2．周辺分布ヒストグラムモデルに基づくGA

本章では，周辺分布ヒストグラムのモデルとして，等幅ヒストグラム（FWH: Fixed width histogram）と等高ヒストグラム（FHH: Fixed height histogram）の二つのモデルについて述べる．

2．1 アルゴリズムの構造

提案アルゴリズムの流れ図を以下に示す．

0. Set the generation counter $t \leftarrow 0$
1. Generate the initial population $P(0)$ randomly.
2. Evaluate functional value of the vectors (individuals) in $P(t)$.
3. Select a set of promising vectors $S(t)$ from $P(t)$ according the functional values.
4. Construct a marginal histogram model $M(t)$ according to the vector values in $S(t)$.
5. Generate a set of new vectors $O(t)$ according to the marginal histogram model $M(t)$.
6. Evaluate the new vectors in $O(t)$.
7. Create a new population $P(t+1)$ by replacing some vectors from $P(t)$ with $O(t)$.
8. Update the generation counter, $t \leftarrow t + 1$.
9. If the termination conditions have not been satisfied, go to step 3.
10. Obtain solutions in $P(t)$.

従来のGAとの違いは交叉オペレータが，モデル構築（ステップ4），モデルに基づく個体の生成（ステップ5）に変わっているだけである．従って，選択オペレータや世代交代モデルには，従来のGAと同様各種の方法を用いることが出来る．

2.2 等幅ヒストグラム
2.2.1 モデルの記述

等幅ヒストグラム(FWH)は、図1に示すような最も直接的な確率モデルである。FWHでは変数 x_j ($j=1, \dots, n$)の探索区間 $[\min_j, \max_j]$ は H 個のビンに分割され、それぞれのビンの推定確率 $P_{FWH}^j[h]$ が次式で行われる。

$$P_{FWH}^j[h] = \left\{ \left| V[i][j] \right| \mid i \in \{0, 1, \dots, N-1\} \wedge h \leq \frac{V[i][j] - \min_j}{\max_j - \min_j} H < h+1 \right\} / N \quad (1)$$

ただし、 $V[i][j]$ は、個体 i の変数 j の値、 N は個体数である。各ビン間の幅 ε は、要求される解の精度によって決まるが、評価関数が多峰性を有する問題では、正しい解を得るために ε の値に以下のような条件が必要である。図2のようにグローバル解が評価値 h_1 、幅 w_1 、局所解が評価値 h_2 、幅 w_2 で近似できる問題を考える。ここで、 $h_1 > h_2$ 、 $w_1 \ll w_2$ と仮定している。このような条件では、世代0においてランダムに生成される個体のローカル解に位置するビンLにおける平均評価値は、 h_2 となる。また同様にグローバル解に位置するビンGにおける平均評価値は $h_1 \times w_1 / (2\varepsilon)$ である。このとき、 $h_2 > h_1 \times w_1 / (2\varepsilon)$ の条件の時には選択オペレータは、ビンLに位置する個体をビンGに位置する個体よりも多く選択する。この結果、ビンLの確率密度の方が、ビンGの確率密度よりも大きくなる。このヒストグラムに基づいて次世代のサンプリングが行われると、ビンLにはビンGよりも多くの個体がサンプリングされる。このようにして集団は局所解に収束していく。これを避けるためにはビンの幅 ε について

$$\varepsilon < \frac{w_1 \times h_1}{2 \times h_2} \quad (2)$$

の条件が必要である。

2.2.2 サンプリング方法

FWHでは、新しい個体は以下のようにして生成される。まず式(1)で与えられる確率に基づいてビンがサンプリングされる。次に新しい個体は、このビン内で一様サンプリングにより変数の値が得られる。これを必要な個体数が得られるまで繰り返される。ビンをサンプリングする最も簡単な方法はルーレット(RW; roulette wheel)を用いる方法である。しかしRWは、統計的サンプリング誤差が大きい。Sをサンプリング総数とすると、ビン h が選ばれる個数Rの密度関数は、

$$P_S^{hj}(R) = \binom{S}{R} \times (P_{FWH}^j[h])^R (1 - P_{FWH}^j[h])^{S-R}, R = 0, 1, \dots, S \quad (3)$$

となる。ただし、 $P_{FWH}^j[h]$ は、ビン h の確率密度である。RをSで正規化した値 $r(r=[0, 1])$ の分布は、

$$\overline{P_S^{hj}}(r) = S \times P_S^{hj}(R/S) \quad (4)$$

として得られる。図3は、サンプリング誤差の例として $P_{FWH}^j[h]=0.1$ と仮定した場合の r の分布を示したものである。

統計的サンプリング誤差の少ない方法に、BakerのSUS (stochastic universal sampling)法がある[Baker 87]。SUSは、シンプルGAにおける選択方法として提案された手法であるが、ここでは、SUSをFWHのサンプリングの方法に拡張し、E-SUS (extended SUS)と呼ぶ。FWHにおけるE-SUSの疑似コードを図4に示す。ここで13行の変数"expected"は、パラ

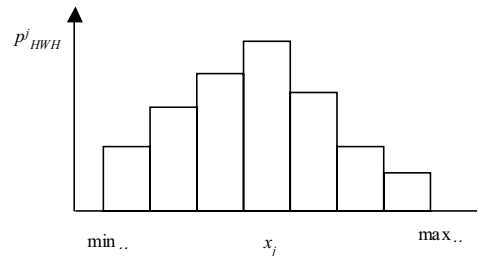


図1 周辺等幅ヒストグラム(FWH)

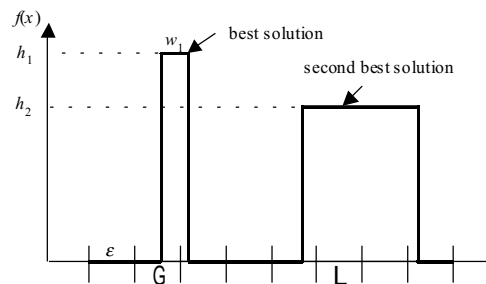


図2 多峰性関数

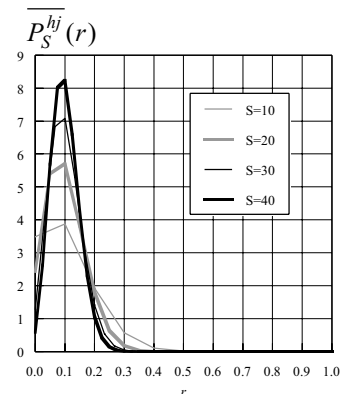


図3 サンプリング誤差

メータ x_j のビン h のサンプル数の期待値である。例えば、あるビンの確率密度が0.155、サンプリング総数が100の場合を考えるとexpected値は15.5である。この場合、E-SUSでは、15個を固定的にサンプリングし、残りの0.5を確率的にサンプリングする。変数間のランダムな組み合わせを得るために、各パラメータ毎にサンプリングシーケンスはランダムにシャッフルされる(8行参照)。3章の実験では、RWとE-SUSの比較も行う。

```

1 //S:sampling individual number, n: number of variables, H: number of bins
2 // p[j][h]: probability density of bin h of variable x_j
3 // t[j][h]: left edge position of bin h of variable x_j
4 // Rand(): generate random number from [0, 1.0]
5 V[S][n]; // array for sampled vectors
6 xh[S]={0,1,2,...,S-1}; // array for random permutation of bin position
7 for(int j=0; j<n; j++){
8     Shuffle(xh); // get a permutation xh[] by shuffling for each parameter
9     double ptr = Rand();
10    double sum = 0.0;
11    int k = 0;
12    for(int h=0; h<H; h++){
13        double expected = p[j][h]*S;
14        for(sum += expected; sum>ptr; ptr++){
15            V[xh[k++]][j] = t[j][h]+(t[j][h+1]-t[j][h])*Rand();
16        }
17    }

```

2.3 等高ヒストグラム(FHH)
2.3.1 モデルの記述

等高ヒストグラム(FHH)では、各ビンの確率密度が $1/H$ の値を持つ。すなわち、各ビンが同数の個体を持つようにモデル化される。FHHの大きな特徴は、各ビンによってその幅が異なり、個体密度の高い領域ではビン幅が動的に縮小することである。この結果、有望な探索領域では、ビン幅が小さくなり高精度な探索が行われる。図5 (a)は、集団の分布が区間[-5,5]において図5 (b)に示すような正規分布 $N(0,1)$ となる場合のFHHの例である。この例では20個のビンが使われており、 $x_j=0.0$ 近辺での幅が狭くなっているため、この近辺からより多くの個体がサンプリングされる。

図4 FWHにおけるE-SUSの擬似コード

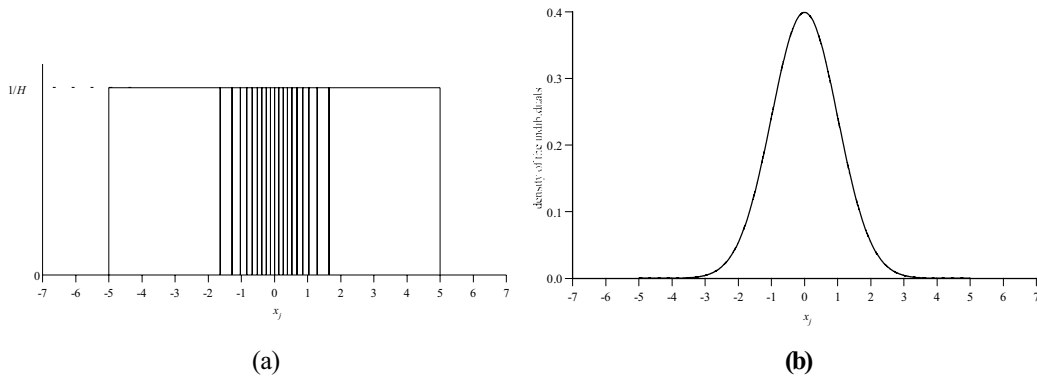


図5 周辺等高ヒストグラム(FHH) (a) FHH, (b) 対応する密度関数

2.3.2 サンプリング方法

FHHのサンプリング方法は、確率密度が全てのビンにおいて同一値 $1/H$ となる以外はFWHのサンプリング方法と基本的に同じである。FHHにE-SUSを適用する場合には、FWHの図4の13行が

```
"double expected = (double)S/H;"
```

となる。

3. 実験

3.1 実験条件

世代交代モデルは、 $(\mu+\lambda)$ -ESと類似のモデルを用いた。まず、サイズ N の集団 $P(t)$ から、ヒストグラムモデルを生成し、そのモデルを基に $N*K$ 個の子個体を生成し評価する。 $P(t)$ の N 個の個体と生成された $N*K$ 個の子個体とからベスト M 個の個体を選択し、次世代の集団 $P(t+1)$ とする。本実験では $K=1$ とした。

本実験に用いたテスト関数は表1に示す通りである。関数 $F_{TwoPeaks}$ は、サブ関数の和からなる関数である。各サブ関数は、図6に示すように、広くて低いピークと高く狭いピークとからなっており、バイナリコーディングGAの研究で良く用いられる騙し関数を連続領域のに拡張したものである。この関数は変数間の絡みを持たない。Rastrigin関数 $F_{Rastrigin}$ も変数間の絡みを持たない。Griewank関数 $F_{Griewank}$ は、変数間に弱い絡みを持つ。Schfewel関数 $F_{Schfewel}$ は、変数間に中程度の絡みを持つ。

ヒストグラムのビン数(H)に関しては、初期状態におけるビン幅が0.1となるように設定した。例えば関数 $F_{TwoPeaks}$ では、ビン数は $12.0/0.1=120$ となる。

3.2 結果

ここでは、(1)解精度を0.1としたときのFWHとFHHの探索特性、(2)FHHの高精度探索特性、(3)FHHのパラメータ次元を増やした場合のスケールアップ特性、の三つの実験結果について述べる。

3.2.1 解精度を0.1としたときのFWHとFHHの探索特性

FWHの探索精度は高々0.1であるので、0.1の精度での探索能力をFWHとFHHで比較した。この結果を表2に示す。サンプリング方法にはRWとE-SUSの両方を用いた。また比較のために実数値GAでよく用いられるBLX-α [Eshelman 93]の結果も示した。

まず、RWとE-SUS二つのサンプリング方法の結果の差を見ると、殆どどの実験でE-SUSが優れていることが分かる。例えば、関数 $F_{TwoPeaks}$ では、RWを用いたFWH(FWH/RW)は、集団サイズ600で20回の実験中最適解を20回発見し、そのとき必要とした関数評価回数の平均(MNE: mean number of evaluations)は、14,620.9である。これに対して、E-SUSを用いたFWH(FWH/E-SUS)では、最適解発見回数は20と同じであるが、それに必要とした集団サイズは300、MNEは7,178.3であり、それぞれFWH/RWの約半分である。

次に、FWHとFHHの性能を比べてみると、FHHの方がいずれのサンプリング方法を用いてもFWHよりも優れている。例えば、関数 $F_{TwoPeaks}$ では、FHH/RWが集団サイズ300最適解を20回発見し、そのときのMNEが8,321.8であるが、FWH/RWは、最適解を20回発見するのに集団サイズ600を必要とし、またそのときのMNEは14,620.9である。

最後にBLX-αとの比較で見ると、関数 $F_{Griewank}$ および $F_{Schwefel}$ では、BLX-αは比較的よい性能を示しているが、関数 $F_{Rastrigin}$ および $F_{TwoPeaks}$ ではFWH、FHHの方が優れている。特に関数 $F_{TwoPeaks}$ では、BLX-αは、20回の実験中一度も最適解を発見することが出来なかった。

3.2.2 FHHの高精度探索特性

2.3節図5に示したようにFHHは、個体分布の密度が高いところは、ピンの幅が縮小し、この結果高精度に解を探索できる特徴を持っている。図7は、関数 $F_{Rastrigin}$ における変数 x_1 でのピン位置の変化の一例を示したものである。世代が進むに従って、ピン位置が最適解の近辺に収束していく状況が分かる。なお、この例ではFHH/RWを用い、集団サイズは800とした。

図8に各関数におけるFHHの収束過程を示す。縦軸は、関数値の集団における平均値である。集団サイズは各関数とも800とした。実験回数は10とし、各実験での関数評価

表1 テスト関数

Function	Domain	Interactions*	H**
$F_{Two-peaks} = n \times 5 - \sum_{i=1}^n f_i$	$[0, 12]^n, n = 20$	non	120
$F_{Rastrigin} = n \times 10 + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i))$	$[-5, 5]^n, n = 20$	non	100
$F_{Griewank} = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n (\cos(x_i / \sqrt{i}))$	$[-5, 5]^n, n = 10$	weak	100
$F_{Schwefel} = \sum_{i=2}^n [(x_i - x_i^2)^2 - (x_i - 1)^2]$	$[-2, 2]^n, n = 5$	medium	40

* Interactions among variables ** Number of bins in each variable

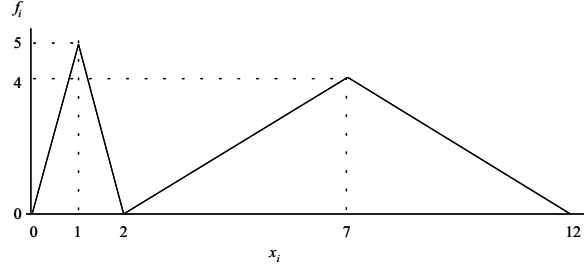


図6 $F_{TwoPeaks}$ のサブ関数

表2 解精度を0.1としたときのFWHとFHHの探索特性

Model	Function	n	Population size											
			100		200		300		400		600		800	
			#OPT	MNE	#OPT	MNE	#OPT	MNE	#OPT	MNE	#OPT	MNE	#OPT	MNE
FWH/RW	$F_{Two-peaks}$	20	0	-	1	4,938.0	7	7,852.4	16	9,854.1	20	14,620.9	20	18,808.0
	$F_{Rastrigin}$	20	0	-	1	7,851.0	6	9,204.0	12	13,154.2	20	19,396.6	20	24,762.7
	$F_{Griewank}$	10	0	-	1	6,527.0	4	9,515.8	14	11,908.6	16	17,884.3	20	23,198.2
	$F_{Schwefel}$	5	13	1,188.9	16	1,927.4	19	3,175.5	20	3,900.2	20	5,282.0	20	6,828.4
FWH/E-SUS	$F_{Two-peaks}$	20	1	2,562.0	14	4,926.5	20	7,178.3	19	9,609.8	20	14,016.6	20	18,537.9
	$F_{Rastrigin}$	20	1	3,478.0	13	6,770.6	20	9,720.2	20	12,616.4	20	18,518.5	20	24,113.5
	$F_{Griewank}$	10	2	2,779.0	7	5,783.3	15	9,220.9	19	12,439.8	19	17,517.2	19	22,502.3
	$F_{Schwefel}$	5	16	1,342.6	19	1,962.3	19	3,302.9	20	4,012.7	20	5,924.8	20	7,597.9
FHH/RW	$F_{Two-peaks}$	20	20	17,952.4	20	9,443.2	20	8,321.8	20	10,305.9	20	14,950.5	20	19,442.9
	$F_{Rastrigin}$	20	20	50,099.8	20	46,139.8	20	21,933.1	20	17,566.2	20	22,847.3	20	28,990.9
	$F_{Griewank}$	10	17	5,445.1	19	6,561.2	18	9,521.2	16	14,076.5	20	17,233.4	19	22,909.5
	$F_{Schwefel}$	5	20	2,758.9	20	4,644.6	20	3,821.2	20	5,969.8	20	6,257.1	20	7,101.6
FHH/E-SUS	$F_{Two-peaks}$	20	20	9,800.4	20	5,405.8	20	7,679.9	20	9,982.4	20	14,380.3	20	19,180.1
	$F_{Rastrigin}$	20	20	16,515.7	20	8,004.2	20	10,177.6	20	12,690.1	20	18,177.4	20	24,879.6
	$F_{Griewank}$	10	18	3,187.4	18	5,680.6	20	8,199.6	20	10,962.5	20	15,782.6	20	20,614.3
	$F_{Schwefel}$	5	20	3,526.7	20	2,994.3	20	3,483.3	20	5,020.2	20	6,234.8	20	8,829.2
BLX-α	$F_{Two-peaks}$	20	0	-	0	-	0	-	0	-	0	-	0	-
	$F_{Rastrigin}$	20	0	-	1	28,521.0	4	47,959.8	11	62,776.0	17	94,994.9	20	124,825.9
	$F_{Griewank}$	10	20	3,316.3	20	6,347.7	20	9,037.2	20	11,947.3	20	17,483.0	20	23,209.9
	$F_{Schwefel}$	5	20	3,305.3	20	3,861.3	20	5,397.8	20	7,141.0	20	10,158.8	20	12,688.0

回数は最大100,000とした。

サンプリング方法の比較ではE-SUSが優れていることが分かる。FHH/RWでは、関数 $F_{TwoPeaks}$ 、 $F_{Rastrigin}$ および $F_{Griewank}$ において数回ローカル解にトラップされている。これに対して、FHH/E-SUSでは、これらの関数において、関数 $F_{Rastrigin}$ で10回の実験中1回だけローカル解にトラップされているのみである。

関数 $F_{Schfewel}$ では、FHHはグローバル解への収束に失敗している。関数 $F_{Schfewel}$ は、変数間に中程度の絡みを持っている。このように周辺分布を用いるFHHは、中程度の強度の変数間の絡みをもつ関数では、精度よく解を得ることに失敗する。

BLX- α オペレータは、関数 $F_{Rastrigin}$ および $F_{Griewank}$ では、FHH/E-SUSと同様な探索特性を持っている。しかし関数 $F_{TwoPeaks}$ では、表2で見たのと同様、グローバル解には収束できなかった。

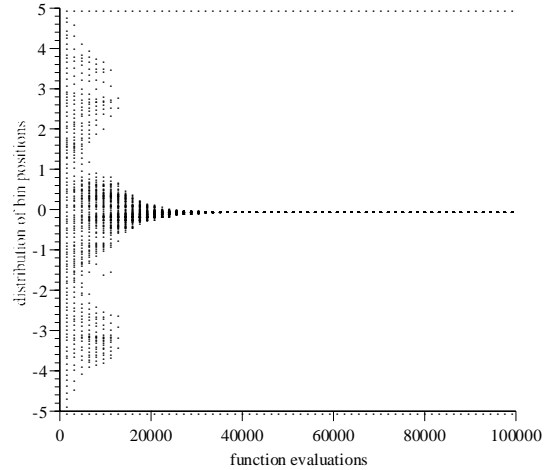


図7 関数 $F_{Griewank}$ のパラメータ x_1 におけるFHHのビン位置の変化の例

3.2.3 FHHのパラメータ数を増やした場合のスケールアップ特性

ここでは、関数 $F_{TwoPeaks}$ のパラメータ数 n を10から100まで増加させたときのスケールアップ特性の結果を見る。ここでの世代交代モデルにはMGG[Satoh 96]を用いた。この関数の最適解は(1.0,1.0,...,1.0)であるので、ベスト個体のパラメータ値が

$$\forall_{i=1,\dots,n} 1.0-0.005 \leq x_i < 1.0+0.005 \tag{6}$$

となったとき、最適解が発見されたと判断する。パラメータ数は10単位で増加させた。このとき10回の実験で10回とも最適解を得るのに必要とした集団サイズと関数評価回数の10回の実験の平均(MNE)の変化を調べた。

図9(a)は、パラメータ数と集団サイズの関係である。集団サイズは、パラメータ数 n に比例($O(n)$)していることが分かる。図9(b)は、MNEの変化である。ほぼ $O(n^{1.56})$ となり、FHHはサブ解の高いミキシング能力と優れたスケールアップ特性を持っていることが明らかになった。

4. 今後の課題

3章で見たように、周辺分布を用いるヒストグラムモデルは、パラメータ間に絡みがない、あるいは弱い絡みの問題では優れた性能を持っている。しかしパラメータ間に中程度あるいは強い絡みを持つ問題ではよい性能が得られない。

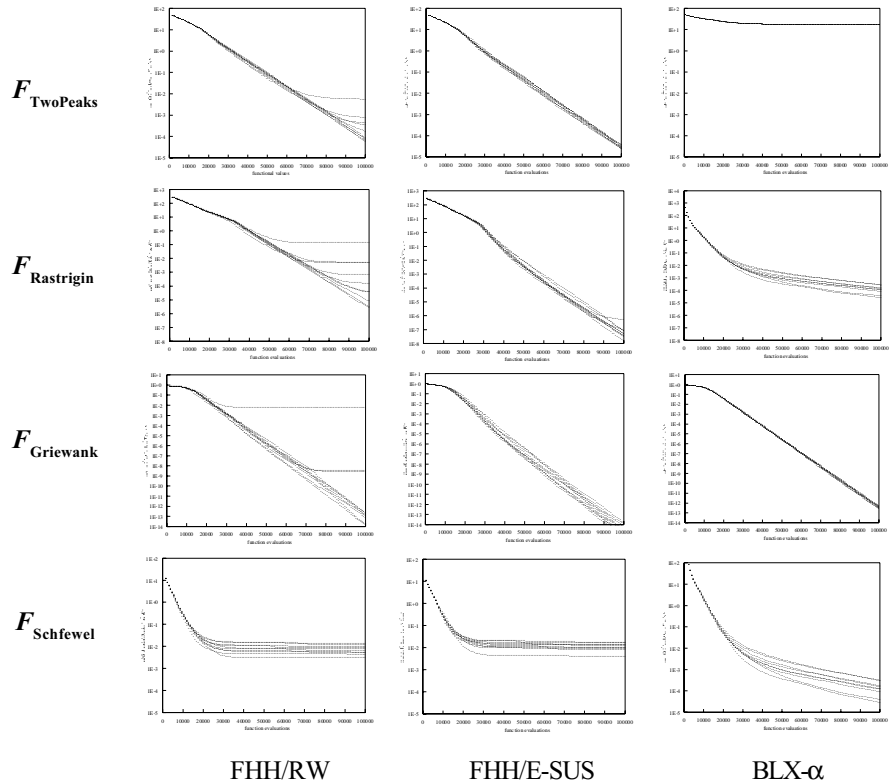
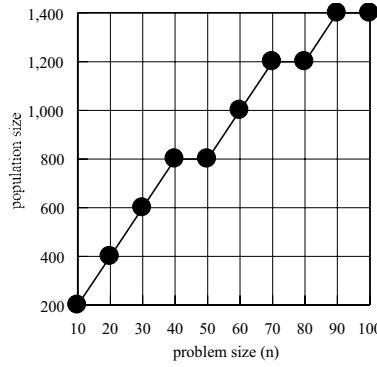
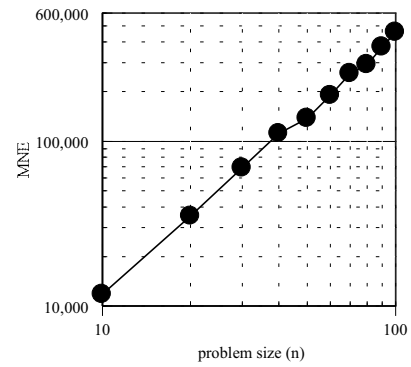


図8 FHHの探索特性

周辺分布を用いるヒストグラムモデルの自然な拡張は、多変量ヒストグラムを用いる方法である。しかし単純な拡張では、総ビン数がパラメータ数 n に対して指数オーダー、すなわち $O(H^n)$ となるので実際的でない[Bosman 99]。従って、ビン数が指数的に増加しない多変量ヒストグラムモデルを構築する工夫が必要となる。図10は、そのようなFHHの一例として「繰り返しFHH」を示したものである。図ではまず x_1 軸で、個体数が同じになるように領域を2分割する。次に x_2 軸で、各々領域を個体数が同じになるように再度領域を2分割する。この分割を、細分割された各々の領域に属する個体数がある決められた個数 K_{\min} になるまで繰り返す。このようにして得られた各々の領域は、周辺分布FHHの各ビンに相当する。子個体は、各分割された領域を等確率でサンプリングし、次にこの領域内で一様サンプリングすることにより生成される。図11は、パラメータ間に強い絡みを持つRosenbrock関数 $F_{\text{Rosenbrock}}$



(a) 集団サイズ



(b) 関数評価回数

図9 スケールアップ特性

における周辺分布FHHと繰り返しFHHの結果である。サンプリング法はともにRWを用いた。集団サイズは3,000とし、 $K_{\min}=2$ とした。この図から、繰り返しFHHは有効な多変量ヒストグラムモデルの一つの候補であることが分かる。さらに一般的なモデルは今後の検討課題である。

$$F_{\text{Rosenbrock}} = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2, \quad -2.048 \leq x_1, x_2 < 2.047 \quad (7)$$

3.2.3項で見たようにFHHは、優れたスケールアップ特性を持っている。これはFHHを用いる確率モデルGAは、サブ解のミキシング能力に優れた性質を持っていることを示している。一方、多変量正規分布混合モデルを用いる確率モデルGAは、変数間にリンクを有する問題に対して優れた性質を有している [Bosman 00] が、ミキシング能力については十分でない。従って、FHHモデルと多変量正規分布混合モデルとを組み合わせる確率モデルGAがこの問題に対する一つの解を与えるものと考えられる。これについても今後の重要な検討課題である。

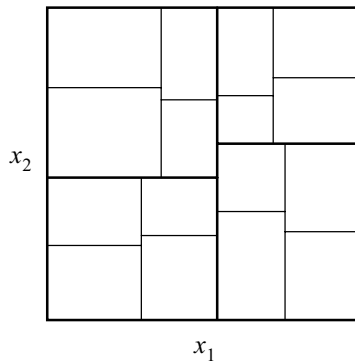


図10 繰り返しFHH

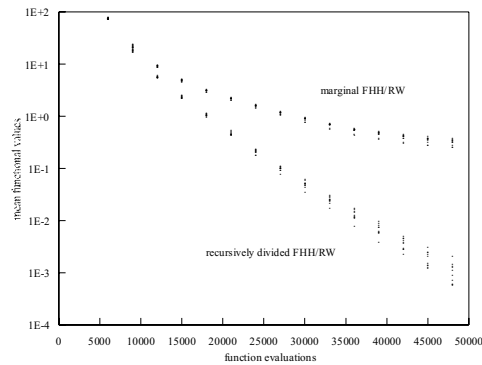


図11 関数 $F_{\text{Rosenbrock}}$ での結果

5. むすび

本論文では、周辺分布を用いるヒストグラム確率モデルGAを提案した。ヒストグラムモデルとして等幅ヒストグラム (FWH) 及び等高ヒストグラム (FHH) を用いた。収束特性ではFHHが、解精度、スケールアップ特性で優れている。

しかし周辺分布を用いるヒストグラム確率モデルGAは、パラメータ間に絡みを有する問題に対しては満足な性能が得られない。これを解決する方法の検討が今後に残された研究課題であり、その考察も行った。

謝辞

本研究の一部は、文部科学省科学研究補助金（基盤研究C）課題番号13680469および同志社大学学術フロンティア「知能情報科学とその応用」研究プロジェクトの一環として行われた。

Professor Goldberg's contribution to the work was sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant F49620-00-0163. Research funding for this work was also provided by a grant from the National Science Foundation under grant DMI-9908252. Support was also provided by a grant from the U. S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0003.

参考文献

- [Baker 87] Baker J. E.: Reducing bias and inefficiency in the selection algorithm, *Proc. of the 2nd International Conference on Genetic Algorithms*, pp. 14-21 (1987).
- [Baluja 94] Baluja, S.: Population-based incremental learning: A method for interacting genetic search based function optimization and coemotive learning, *Tech. Rep. No. CMU-CS-94-163*, Carnegie Mellon University (1994).
- [Baluja 97] Baluja, S. and Davies: Using optimal dependency-trees for combinatorial optimization: learning the structure of the search space, *Tech. Rep. No. CMU-CS-97-107*, Carnegie Mellon University (1997)
- [Bosman 99] Bosman, P. and Thierens, D.: An algorithmic framework for density estimation based evolutionary algorithms, *Tech. Rep. No. UU-CS-1999-46*, Utrecht University (1999).
- [Bosman 00] Bosman, P. and Thierens, D.: Mixed IDEAs, *Tech. Rep. No. UU-CS-2000-45*, Utrecht University (2000).
- [De Bonet 97] De Bonet, J. S., Isbell, C. L. and Viola, P.: MIMIC: Finding optima by estimating probability densities, In Mozer, M. C., Jordan, M. I., and Petsche, T. (Eds): *Advances in neural information processing systems*, Vol. 9, pp. 424-431. (1997).
- [Eshelman 93] Eshelman, L. J. and Shaffer, J. D.: Real-coded genetic algorithms and interval-schemata, *Foundation of Genetic Algorithms 2*, Morgan Kaufmann, pp. 187-202 (1993).
- [Harik 98] Harik, G., Lobo, F. G., and Goldberg, D. E.: The compact genetic algorithm, *Proc. of the International Conference on Evolutionary Computation 1998 (ICEC 98)*, pp. 523-528 (1998).
- [Harik 99] Harik, G.: Linkage learning via probabilistic modeling in the ECGA, *IlligAL Technical Report 99010*, University of Illinois at Urbana-Champaign, Urbana, Illinois (1999).
- [Larranaga 99] Larranaga, P., Etxeberria, R., Lozano, J.A., and Pena, J.M.: Optimization by learning and simulation of bayesian and gaussian networks, *University of the Basque Country Technical Report EHU-KZAAIK -4/99* (1999).
- [Mühlenbein 96] Mühlenbein, H and Paaß, G.: From recombination of genes to the estimation of distribution I. Binary parameters, *Proc. of the Parallel Problem Solving from Nature - PPSN IV*, pp. 178-187 (1996).
- [Pelikan 99] Pelikan, M., Goldberg, D. E., and Cantu-Paz, E.: BOA: The Bayesian optimization algorithm, *Proc. of the Genetic and Evolutionary Computation Conference 1999 (GECCO-99)*, Morgan Kaufmann, San Francisco, CA (1999).
- [Pelikan 00a] Pelikan, M., Goldberg, D. E., and Cantu-Paz, E.: Linkage problems, distribution estimate, and Bayesian network, *Evolutionary Computation*, Vol. 8, No. 3, pp. 311-340 (2000).
- [Pelikan 00b] Pelikan, M., Goldberg, D. E., and Lobo, F. G. : A survey of optimization by building and using probabilistic models, *Computational Optimization and Applications*, Kluwer Academic Publishers (in press).
- [Sebag 98] Sebag, M. and Ducoulombier, A.: Extending population-based incremental learning to continuous search spaces, *Proc. of the Parallel Problem Solving from Nature - PPSN V*, pp. 418-427 (1998).
- [Servet 97] Servet, I. L., Trave-Massuyes, L., and Stern, D.: Telephone network traffic overloading diagnosis and evolutionary computation techniques, *Proc. of the Third European Conference on Artificial Evolution (AE 97)*, pp. 137-144 (1997).
- [Tsutsui 01] Tsutsui, S., Pelikan, M., and Goldberg, D. E.: Evolutionary algorithm using marginal histogram models in continuous domain, *Proc. of the Optimization by Building and Using Probabilistic Models (OBUPM) 2001, GECCO-2001 Birds-of-a-feather Workshops* (2001).