ベイジアンネットワークを遺伝的操作に利用した実数値遺伝的アルゴリズム

同志社大院 ○吉田 純一　同志社大工 廣安 知之　同志社大工 三木 光範

Real-Coded Gentic Algorithm using Bayesian Network as Genetic Operator

○ Jun-ichi YOSHIDA, Graduate School of Engineering, Doshisha University
Tomoyuki HIROYASU, Faculty of Engineering, Doshisha University
Mitsunori MIKI, Faculty of Engineering, Doshisha University

Abstract:
In this paper, we propose, Gaussian Optimization Algorithm (GOA), which generate a new search point by using gaussian network and estimation of distribution. Algorithms where offsprings (new search points) are generated according to estimated probability model of the parents are called Estimation of Distribution Algorithms (DEAs) or Probabilistic Model-Building GAs (PMBGAs). The proposed GOA is one of the DEAs. The GOA can apply for continuous optimization problems and elements of gaussian network are design variables. In the GOA, by using Q values, graph structure of gaussian network is chosen probabilistically from the candidate networks. Q Learning decides these Q values of the candidate networks and these values are reinforced during the searching. Through the numerical examples, the effectiveness of GOA is examined with standard test functions. From the results, it is made clarified that GOA can find a good solution with smaller calculation costs, especially in the test function that has epstasis between the design variables. In this paper, the gaussian network which is derived from the GOA is discussed. Analyzing the derived network, some important design variables are determined.

1 はじめに

一般的な遺伝的アルゴリズム（Genetic Algorithms: GA）[1, 2] では新しい探索点（個体）を生成するために交叉，突然変異が用いられる．交叉には親個体のもつ良好な部分解を交換するという利点があるが，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 [3]．近年，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個体の統計的な情報を用いて個体を生成する確率モデル GA に関する研究がさかんに行われている．本研究では，実数値の染色体をもつ GA において，個体の生成を Q-Learning によって学習されたベイジアンネットワーク（ガウシアンネットワーク）に基づいて生成するガウシアン最適化アルゴリズムを提案する．

2 確率モデル GA

2.1 GA の問題点

GA は生物の進化の過程を工学的に模倣した最適化手法であり，さまざまな問題に適用可能であるが，一般に次のような問題点が知られている [2]．

・パラメータ設定が困難
・計算時間の増大
・良い形質の継承が困難
・最適解の保証がない

これらのうち，形質の継承には GA の交叉オペレータの影響が大きい．交叉の役割やその有効性については多くの報告がなされている．交叉は短いサブ問題から構成される問題や，小規模な問題においては効率よくビルディングブロックを構成することが可能だが（ビルディングブロック仮説 [1]），一方で，新しくビルディングブロックを発見するのは稀であり，むしろ親個体のもつビルディングブロックを破壊する場合が多いと報告されている [3]．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（Probabilistic Model-Building Genetic Algorithms: PMBGAs）に関する研究がさかんに行われている [4]．

2.2 分布推定アルゴリズム

確率モデル GA では，従来の GA と同様にランダムに初期母集団を生成する．母集団内から選択オペレータによって有望な個体が選択される．従来の GA が交叉によって新しい個体を生成するのに対し，確率モデル GA では選択された個体群の確率分布の推定が行われ，この分布に従って新しい個体が生成される．この処理を終了条件を

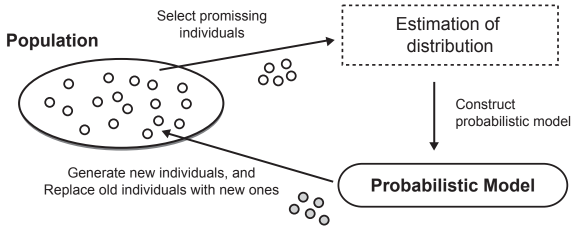

Fig. 1　分布推定アルゴリズムの基本概念

満たすまで繰り返すことで解探索を行う．このような原理に基づいた手法は分布推定アルゴリズム（Estimation of Distribution Algorithms :EDA）とも呼ばれる [5]．EDA の概念図を Fig. 1 に示す．

確率モデル GA は，コーディング法（バイナリストリングまたは実数）とビット間の依存関係を考慮するか否かで分類が可能である．バイナリストリングの染色体を持つモデルの先行研究は多く，さまざまな手法が提案されている [4]．設計変数間の依存関係を考慮しない手法としては，PBIL（Population Based Incremental Learning Algorithm）, UMDA（Univariate Marginal Distribution Algorithm）, cGA（compact GA）がある．これらはいずれも設計変数間に依存関係のない問題においては効率的な探索が期待できる．

設計変数間の依存関係を考慮したものとしては，2 変数の関係を考慮したものに MIMIC（Mutual-Information-Maximization Input Clustering）がある．また多変量の依存関係を考慮した手法に，ECGA（Extended Compact GA）や確率モデルにベイジアンネットワークを用いた BOA（Bayesian Optimization Algorithm）[6] がある．これらの手法は多変量の依存関係の解析と確率モデルの構築に計算量が必要となるが，変数間に複雑な依存関係が存在する問題においても有効に機能すると期待できる．

一方，実数値の染色体を持つ確率モデル GA の研究は少ない．そこで，本論文では実数値の確率モデル GA においてベイジアンネットワークを用いて個体を生成する手法を提案する．ただし，一般にベイジアンネットワークでは離散値を扱うことが多いが，連続変数を扱うためには確率密度関数として正規分布を利用することがある．このようなベイジアンネットワークを特にガウシアンネットワーク [7] と呼ぶ．そこで，本論文では提案する手法をガウシアン最適化アルゴリズム（Gaussian Optimization Algorithm: GOA）と呼ぶことにする．GOA においては，ネットワークのグラフ構造は Q-Learnig によって学習される．

3　ベイジアンネットワーク

ここでは，GOA において個体の生成に用いられるベイジアンネットワークについて簡単に述べる．一般に，確定値を得ることが難しい問題においては，これらを体系的に取り扱うために確率変数を用いることができる．確率変数をノードで表し，因果関係や相関関係といった依存する関係を持つ変数の間にリンクを張ったグラフ構造による確率モデルが確率ネットワークと呼ばれ，その中でとくにリンクが因果関係の方向に向きを持ち，このリンクをたどったパスが循環しない，非循環有向グラフで表されるモデルがベイジアンネットワークである [8]．

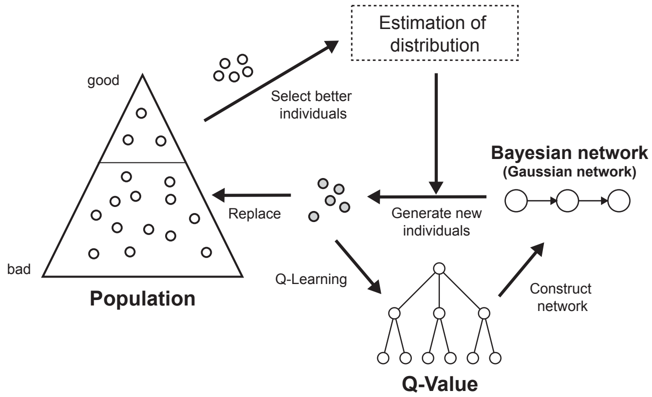

Fig. 2　ガウシアン最適化アルゴリズムの概念図

ベイジアンネットワークは確率変数間の定性的な依存関係をグラフ構造によって表し，変数間の定量的な依存関係はその変数の間に定義される条件付き確率によって表すことで問題領域をモデル化する．

変数 X_i と X_j があり，「if $X_i = a$ then $X_j = b$」というルールが成立しているとき X_j が X_i に依存しているという．しかし，このようなルールが常に成立するとは限らない．こうした不確実性を吸収し，また依存関係の程度を定量的に表すため「$X_i = a_i$ であるとき $X_j = b$ である確率」のような条件つき確率を考えることができる．このときベイジアンネットでは向きのついたリンクによって $X_i \rightarrow X_j$ と表し，この場合の依存関係は条件付確率 $P(X_j|\pi(X_j))$ で定義される．

一般的にベイジアンネットでは離散確率変数を扱うが，連続確率変数でベイジアンネットワークを構成することもできる．この場合には条件付き確率は連続的な関数によって表す必要があり，適当なパラメータを導入して特定の関数族を仮定することで表現する．とくにガウス分布が用いられることが多く，このように構成したベイジアンネットワークはガウシアンネットワーク [7] と呼ばれる．

4　ガウシアン最適化アルゴリズム（GOA）

4.1　アルゴリズム概説

提案するガウシアン最適化アルゴリズム（GOA）のアルゴリズムは以下のようなものである．また，GOA の概念図を Fig. 2 に示す．

1. **初期個体の生成：**
   世代数 $t \leftarrow 0$
   ランダムに初期母集団 P(0) を生成する．

2. **選択：**
   世代 t の母集団 P(t) から，適合度の高い個体群 S(t) を抽出する．

3. **分布推定：**
   S(t) の分布を推定する．

4. **ネットワークの構築：**
   Q 値に基づきネットワーク N の構造を決定する．

5. **個体の生成：**
   ネットワーク N にエンコードされた結合分布（joint distribution）にもとづいて新しい個体 O(t) を生成する．

6. **ネットワークの学習：**
   O(t) を評価．O(t) のうち S(t) の平均値よりも高い平均値を持つ個体を生成したネットワークに報酬を与える．

7. P(t) の個体の一部を O(t) と置き換えることで次世代の母集団 P(t+1) を構築する．
   世代数 t ← t+1.

8. 2 に戻り，終了条件を満たすまで繰り返す．

以下でそれぞれの操作について簡単に説明する．

初期化，サンプル個体群の選択　まずランダムに初期母集団が形成され，次に現在の母集団から良好な個体（サンプル個体群）が選択される．このとき，母集団サイズに対するサンプル個体群の割合を規定するサンプル率 s を設定する必要がある．選択には任意の選択手法を用いることができるが，ここでは単に適合度の上位個体をサンプル個体群として採用する．

分布の推定　選択したサンプル個体群の分布を推定するために，各変数ごとの平均 $\bar{x}$ と標準偏差 σ_x，および各設計変数間の相関係数 $\rho_{x_1 x_2}$（以後 ρ）を求める．これらの値をもとに各設計変数は正規分布すると推定する（Fig. 3）．

ネットワークの構築　提案する GOA ではネットワークの構造の学習に Q-Learning を用いる．個体を生成するためのネットワークの構造は，Q 値によって 1 個体ごとに確率的に決定される．Q-Learning によるネットワークの構造の決定と学習については 5 節で詳しく述べる．

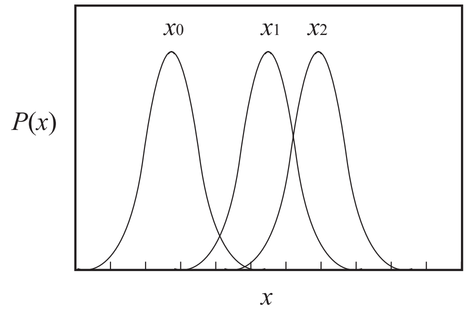

Fig. 3　各変数の正規分布の例

個体の生成　ネットワークを用いて次世代の個体を構築する．ネットワークが $x_0 \rightarrow x_1 \rightarrow x_2$ という関係を表すとき，生成される個体の設計変数は次のように行われる．

まず，ネットワークの始点 x_0 の値が，分布の推定で求めた $\bar{x_0}$ と σ_{x_0} によって定義される正規分布に従って生成される．次に，x_1 の値を，x_0 の値をもとに式 (1) に示す２変量正規分布を満たすように決定する．

$$f(x,y) = \frac{1}{(\sqrt{2\pi})^2\sqrt{\sigma_x\sigma_y - \rho^2}} \cdot e^{-\frac{1}{2}\cdot\frac{1}{\sigma_x\sigma_y-\rho^2}\{\sigma_y(x-\bar{x})^2-2\rho(x-\bar{x})(y-\bar{y})+\sigma_x(y-\bar{y})^2\}} \tag{1}$$

個体の評価とネットワークの学習　生成された個体の評価値を求める．評価値が良好であればこの個体を生成したネットワークに報酬を与え，ネットワークの Q 値を更新する．ここでは，評価値がサンプル個体群の平均よりも高い場合に報酬を与える．

個体の置き換え　このようにして生成された個体をサンプル個体群以外の個体と置き換えることで，次世代の母集団を構成する．

以上の操作を終了条件を満たすまで繰り返すことで，GOA による解探索が行われる．

5　ネットワークの学習

一般に，ベイジアンネットワークの学習およびその利用には,（1）確率変数の選択,（2）グラフ構造の決定,（3）条件付き確率の推定,（4）確率推論の実行という４つのステップがある．ベイジアンネットワークを利用するためにそれぞれについて検討する必要がある．

本論文において（1）は連続関数の各設計変数である．（3）の条件付き確率の選定はサンプル個体群から得た正規分布に従う.（4）の確率推論の実行法は前節で述べたとおりである．

ネットワークのグラフ構造の決定のために，本論文では強化学習の代表的手法である Q-Learning を用いる．ただし，グラフの構造は，Fig. 4 に示すような 1 入力 1 出力とし，2 設計変数間の依存関係を考慮する．

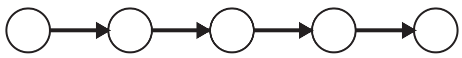

Fig. 4　1 入力 1 出力のネットワーク

5.1　Q-Learning

強化学習法（Reinforcement Learning）は，学習エージェントが未知の環境内で試行錯誤を繰り返すことにより環境に適応した制御法を学習していく手法である [9]．ここでは，代表的な強化学習法である Q-Learning について述べる．

環境内における状態集合を S，エージェントが実行可能な行動の集合を A とする．集合 S と A の要素数はともに有限であるとする．時刻 t で観測された状態 $s_t \in S$ においてエージェントが行動 $a_t \in A$ を実行し，時刻 $t+1$ において状態 s_{t+1} に推移し，報酬 r_{t+1} を得たとすると，状態 s_t の行動価値関数 $Q(s_t, a_t)$ は，次式を用いて更新される．

$$
\begin{aligned}
Q(s_t, a_t) \quad \leftarrow \quad & (1-\alpha)Q(s_t, a_t) \\
& + \alpha[r_t + \gamma \max_a Q(s_{t+1}, a_t)] \qquad (2)
\end{aligned}
$$

ここで，$\alpha (0 < \alpha \leq 1)$ は学習率，$\gamma (0 < \gamma \leq 1)$ は割引率である．行動価値関数とは状態と行動の組に対する評価を見積もる関数であり，Q-learning ではその評価値を Q 値と呼ぶ．エージェントは学習した Q 値の大きさに関係した確率で行動を決定する．ここでは行動選択法として ϵ-greedy 選択を用いる．ϵ-greedy 選択では，ϵ の確率でランダム，それ以外の場合は最大の Q 値をもつ行動が選択される．

5.2　Q-Learning によるネットワークの学習

ガウシアンネットワークが Fig. 4 に示したような 1 入力 1 出力の構造をとるとき，3 変数の問題を扱う場合に存在しうるネットワークは 3!通りである．GOA では Fig. 5 のようなツリー構造でこれを表現し，各パスに Q 値を割り当てる．

個体の生成の際には 1 個体ごとに，Q 値の大きさに関係した確率でネットワークが決定される．生成された個体の評価値がサンプル個体群の平均値よりも良好な場合には，このネットワークに対して報酬が与えられ，Q 値が更新される．いま $x_0 \rightarrow x_1 \rightarrow x_2$ というネットワークから高い評価値をもつ個体が生成されたとすると，このネットワークに報酬が与えられ，式 (2) に従って x_2 と x_1 を結ぶパスの Q 値が更新される．

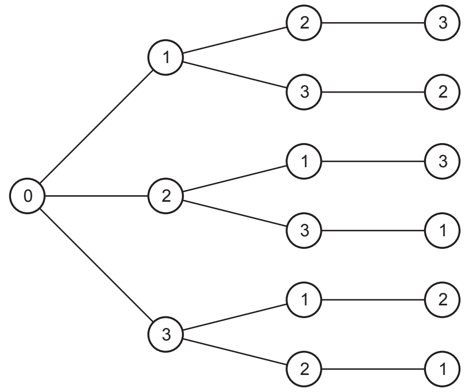

Fig. 5　3 変数からなるネットワークを表現するツリー

探索開始時点では Fig. 5 のツリーのすべてのパスに割り当てられた Q 値は 0 である．どのようなネットワークが選択されるかはランダムである．しかし，適合度の高い個体を生成したネットワークにのみ報酬が与えられるため，各パスの持つ Q 値に差が生じる．結果として，探索が進むにつれ良好なネットワークほど高い確率で選択されるようになり効率的な個体の生成が期待できる．

6　数値実験

GOA の性能を検討するために，連続関数最大化問題において次の 4 つの手法を比較する．

・GOA
・ネットワークなし GOA（各変数ごとに正規乱数を発生）
・実数値 GA（BLX-α）[10]
・実数値 GA（UNDX）[10]

実験に用いた GOA のパラメータは次のとおりである．母集団サイズ 800, サンプル率 0.1，突然変異率 0.1．また，Q-Learning における学習率は 0.1，割引率は 0.3 とし，行動選択法には ϵ-greedy 選択（$\epsilon = 0.5$）を用いた．

比較に用いた実数値 GA のパラメータは母集団サイズ 800, 交叉率 1.0 とし，交叉法として BLX-α（$\alpha = 0.36$）および UNDX（$\alpha = 1.0, \beta = 0.35$）を用いた．

6.1　対象問題

本論文で対象とするのは 5 設計変数の Rastrigin 関数と Rosenbrock 関数である．Rastrigin 関数は格子状に配置された多くの局所最適解をもつ多峰性関数であるが，設計変数間に依存関係はないため GA より解探索が容易な関数として知られている．一方，Rosenbrock 関数はなだらかな単峰性関数であるが，設計変数間に強い依存関係があるため GA により最適解を得るのが難しい関数として知られている．それぞれの式を以下に示す．

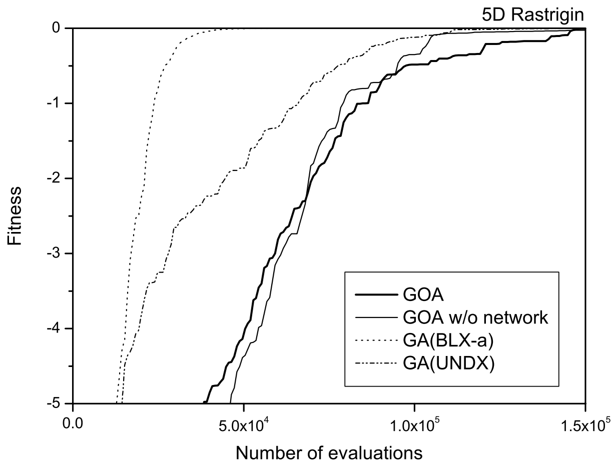

Fig. 6　適合度の推移（Rastrigin）

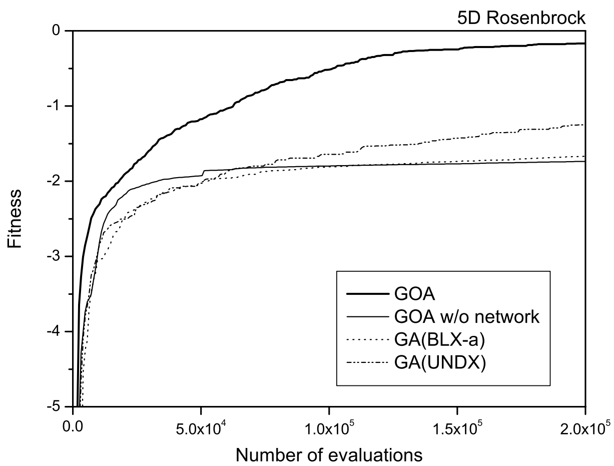

Fig. 7　適合度の推移（Rosenbrock）

$$F_{\mathrm{rastrigin}}(x) = 10n + \sum_{i=1}^{n}\left[{x_i}^2 - 10\cos(2\pi x_i)\right] \quad (3)$$
$$-5.12 < x_i < 5.12, n = 5$$
$$F_{\mathrm{rosen}}(x) = \sum_{i=2}^{n}[100(x_1 - {x_i}^2)^2 + (x_i - 1)^2] \quad (4)$$
$$-2.048 < x_i < 2.048, n = 5$$

6.2　適合度の推移

Rastrigin 関数と Rosenbrock 関数における評価計算回数に対する適合度の推移を Fig. 6, 7 に示す．なお，結果は 20 試行の平均値である．

まず，GOA とネットワークなしの GOA を比較する．Rastrigin 関数においては比較的良く似た傾向を示すのに対し Rosenbrock 関数においては傾向がまったく異なっている．ネットワークなしの GOA では各変数の値はサンプル個体群の正規分布に従って独立に決定されており設計変数間の依存関係については考慮されていない．Rastrigin 関数は設計変数間に依存関係がないため各変数を独立に決定するネットワークなし GOA でも良好な結果が得られる．しかしながら，設計変数間に強い依存関係を有する Rosenbrock 関数では，各変数を独立に探索しても評価値は改善されにくい．このため GOA とネットワークなしの GOA の間におおきな性能差が生じたと考えられる．

次に従来手法である実数値 GA と GOA の性能を比較する．Rastrigin 関数では交叉法に BLX-α を用いた GA が良好な性能を示している．Rosenbrock 関数においては BLX-α やネットワークなし GOA は早い段階で停滞してしまうのに対し，UNDX と GOA は継続的に評価値が改善されている．GOA と UNDX を比較すると GOA のほうがより効率的な探索を行っていることが分かる．

これらの結果から，設計変数間に依存関係のない問題においてはネットワークの学習とそれに基づく個体の生成の効果は少ないが，設計変数間に依存関係のある問題においては GOA は非常に効率的な探索が可能であるといえる．

6.3　ネットワークの学習

6.2 節の結果から，設計変数に依存関係のある問題においてはガウシアンネットワークによる依存関係の学習が有効に機能する可能性が確認された．ここでは，ネットワークの決定に用いられた Q 値のツリーに注目する．Rosenbrock 関数に GOA を適用した際の 100 世代目における Q 値のツリーの一部を Fig. 8 に示す．図中のパスの太さが Q 値の大きさに対応している．Q 値が 0.1 以下の場合は点線で示した．

Fig. 8 中の (a) はネットワークの始点を決定するための Q 値を示している．ここでは，変数 0, 変数 1 のとき Q 値が大きく，それ以外の時は Q 値が小さい．(b), (c) は (a) の子ノードの一つでありそれぞれネットワークの始点が変数 0 の時と変数 4 の時の一部を示している．

Q-Learning では報酬は末端から伝播するため，(a) において Q 値の大きかった変数 0 の子ノードである (b) の Q 値は大きく，逆に，(c) の Q 値は非常に小さい．しかしながら，(c) においても一部のパスは Q 値が大きくなっている．ここでは変数 0,1 をつなぐパスあるいは，それに至るパスの重みが強い傾向がある．このことから，Rosenbrock 関数においては変数 0, 1 の値が評価値に大きな影響を与えていると推測ことができる．式 (4) に示した Rosenbrock 関数の定義式に注目すると x_1(Fig. 8 では変数 0 に相当) はすべての変数との間に依存関係を有していることから，変数 0 の値が評価値に及ぼす影響は大きい．

このように，GOA では学習後の Q 値を検討することにより，対象問題の依存関係に関する知見を得ることもできると考えられる．

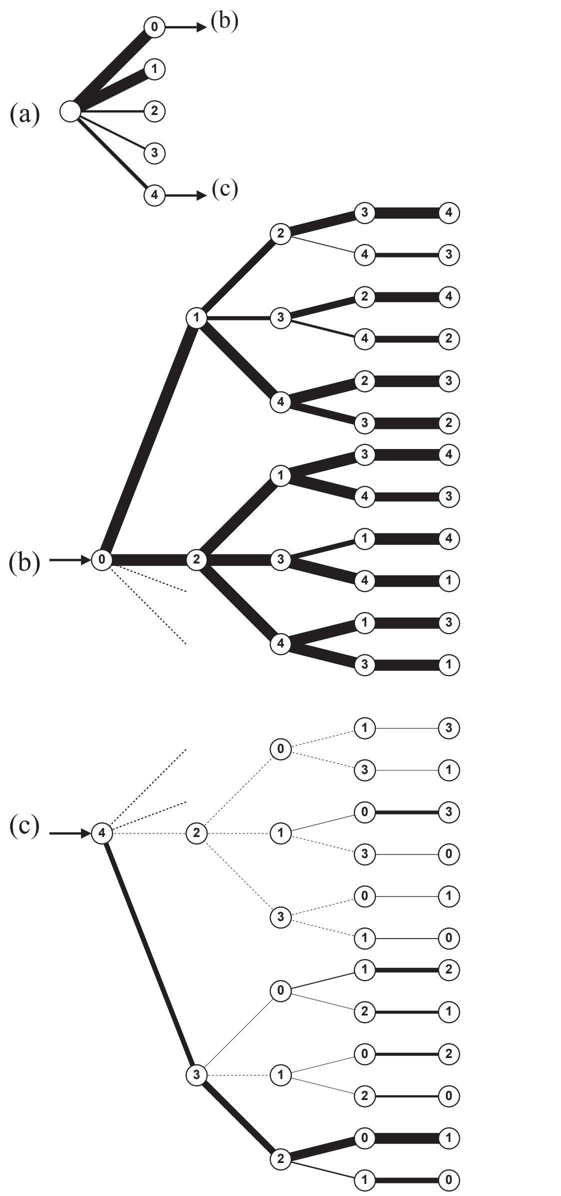

Fig. 8　学習した Q 値（Rosenbrock 関数，100 世代目）

7　おわりに

本論文では，Q-Learning によって学習されたガウシアンネットワークに基づいて個体を生成するガウシアン最適化アルゴリズム（Gaussian Optimization Algorithm: GOA）を提案し，その性能について検討した．数値実験の結果，設計変数間に依存関係のある問題において特に良好な性能を示すことが明らかとなった．また，学習後の Q 値を検討することにより，対象問題の依存関係について検討することができると考えられる．

今回はガウシアンネットワークの有効性を検討するために，シンプルな選択手法や世代交代モデルを用いたが，他の手法を使用することも可能である．また，各種のパラメータについてもチューニングの余地がある．今後はこれらについても検討を行う．

また，その他のテスト関数での性能の検証および，実問題にも GOA を適用し，最適化を行うと同時に学習後のネットワーク構造に基づく対象問題の特性の推測を行う予定である．

謝辞

本研究は文部科学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「知能情報科学とその応用」における研究の一環として行った．ここに記して謝意を表す．

参考文献

[1] J.H.Holland. *Adaptation in Natural and Artificial Systems.* University of Michigan Press, 1975.

[2] D.E.Goldberg. *Genetic Algorithms in Search Optimization and Machine Learning.* Addison-Wesley, 1989.

[3] Annie S. Wu, Robert K. Lindsay, and Rick L. Riolo. Emprical observation on the roles of crossover and mutation. *Proc. 7th International Conference on Genetic Algorithms*, pp. 362–369, 1997.

[4] M. Pelikan, David E. Goldberg, and Fernando Lobo. A survey of optimization by building and using probabilistic models. IlliGAL Report 99018, University of illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, 1998.

[5] H. Muhlenbein and G. Paa. From recombination of genes to the estimation of distributions. *Parallel Problem Solving From Nature IV*, pp. 178–187, 1996.

[6] Martin. Pelikan, David. E. Goldberg, and Erick Cantú-Paz. Boa: The bayesian optimization algorithm. *Proc. Genetic and Evolutionary Computation Conference (GECCO) 1999*, Vol. 1, pp. 525–532, 1999.

[7] Dan Geiger and David Heckerman. Learning gaussian networks. *Proceedings of the 10th Uncertainty in Artificial Intelligence*, 1994.

[8] 本村陽一, 佐藤泰介. ベイジアンネットワーク -不確定性のモデリング技術-. 人工知能学会誌, Vol. 15, No. 4, pp. 575 – 582, 2000.

[9] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning.* MIT Press, 1998.

[10] 小野功, 山村雅幸, 喜多一. 実数値 ga とその応用（解説）. 人工知能学会誌, Vol. 15, No. 2, 2000.

出典：

第 14 回自律分散システム・シンポジウム論文集,
pp. 117-122
(2002 年 1 月 25 日・26 日・東京)

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)