

人工蜂コロニー最適化を用いた Dynamic Bayesian Network の構造学習法
A Structure Learning Method of Dynamic Bayesian Networks
Using an Artificial Bee Colony Optimizer

山室 洋貴[†] 中野 秀洋[‡] 宮内 新[†]
Hiroki Yamamuro Nakano Hidehiro Arata Miyauchi

1. はじめに

Dynamic Bayesian Network (DBN)は確率的に変動する時系列の振る舞いを近似するモデルであり、盛んに研究されている。時系列の背後に存在するダイナミクスが未知である場合、DBN の構造学習が必要である。DBN の評価指標として尤度やネットワークの複雑性などが用いられるが、対象とする時系列の変数が増加すると評価指標の計算量が極めて増大する。また、この構造学習は高次元な多目的最適化問題となるため、効率的な学習アルゴリズムが望まれる。本研究では、高次元問題に対して高い性能をもつ人工蜂コロニー最適化を用いた DBN の構造学習法を提案する。数値実験を行い、提案手法の有効性を示す。

2. ベイジアンネットワーク

ベイジアンネットワーク(Bayesian Network: BN)は複数の変数をノードで表し、その依存関係を方向付きリンクで表した確率モデルである[1]。またノードからリンクをたどって同じノードに循環しない非循環有向グラフで表される。BN は不確実性を含む確率の計算が可能で、金融状況の予測やインターネット広告のマーケティング、医療診断、画像認識などで使用されている。

3. ダイナミックベイジアンネットワーク

ダイナミックベイジアンネットワーク(Dynamic Bayesian Network: DBN)は BN の各変数を時系列について拡張したモデルである[2]。DBN は時刻 0 における各変数にリンクを持つ事前確率 $\Pr(X[0])$ を表す Prior Network DBN_0 と、時刻 T から T + 1 の変数間にリンクする遷移確率 $\Pr(X[t+1] | X[t])$ を表す Transition Network DBN_t によって構成される(図1)。時刻 $t=0, 1, 2, \dots, T$ における各変数の結合確率は式(1)のように表すことができる。

$$DBN_0(X[0]) \cdot \prod_{t=0}^{T-1} DBN_t(X[t+1] | X[t]) \quad (1)$$

図1 DBN

4. ベイジアンネットワークの構造学習

BN の構造が未知の場合、学習データ群からその構造を自動的に求めるための学習を行う。DBN の構造学習には構造を評価する評価基準と、その基準から評価値を得て学習する手法が必要である。評価基準として赤池情報量規準(AIC)やベイズ情報量規準(BIC)などがある[3]。学習

方法としては遺伝的アルゴリズム(GA)、粒子群最適化(PSO)などがある[4][5]。

5. 人工蜂コロニー

人工蜂コロニー(Artificial Bee Colony: ABC)とはミツバチの群れの振る舞いを模倣したアルゴリズムであり、メタヒューリスティックの一つである[6]。粒子群最適化(PSO)や遺伝的アルゴリズム(GA)と比べ局所解に陥りにくく高次元問題に対して効果的に機能するアルゴリズムである。ABC の探索アルゴリズムは3種類のハチに模した 1: employed bee, 2: onlookers, 3: scout の探索に大きく別れている。employed bee は以下の式(2)に示すように自身の周囲のランダムに選んだ位置ベクトル(x_{ij})を使って新位置ベクトル(v_{ij})を決定する。

$$v_{ij} = x_{ij} + \phi(x_{ij} - x_{ij}) \quad (2)$$

i は探索点番号、 j はランダムで選択した次元番号、 ϕ は $[-1, 1]$ の一様乱数、 t は i 以外のランダムで選択した探索点番号である。

onlookers は employed bee の探索結果からルーレット選択によって選ばれた探索点に employed bee と同様の処理を行う。

scout は employed bee, onlookers 探索で一定回数以上新しい解が見つからなかった探索点を破棄しランダムで次の探索点を決定する。

6. 提案手法

6.1 ABCによるBNの構造学習

ABC は高次元問題であっても局所解に陥りにくく解を求められるという利点がある。従来手法では困難であった高次元問題である大規模な BN の構造学習を ABC を用いれば効果的に行うことができる。基本的な ABC は連続値の探索を行う手法であるため離散化する必要がある。そこで式(2)を以下の式のように変更し ABC でのすべての探索点の値を 0 または 1 になるようにする。

$$x_{ij} = x_{ij} \\ \text{if } r > \rho \text{ then } v_{ij} = x_{ij} \\ \text{else } v_{ij} = \overline{x_{ij}} \quad (3)$$

$$x_{ij} \neq x_{ij} \\ \text{if } r > 0.5 \text{ then } v_{ij} = 0 \\ \text{else } v_{ij} = 1 \quad (4)$$

r は $[0,1]$ の一様乱数、 ρ は閾値である。現探索点と新探索点のランダムで選ばれた次元について値が同じであるとき、式(3)のようにある一定の確率で反転する。一方、異なる値

[†] 東京都市大学工学研究科, Graduate School of Engineering, Tokyo City University

[‡] 東京都市大学知識工学部情報科学科

Department of Computer Science, Faculty of Knowledge Engineering, Tokyo City University

である場合は 0 か 1 のランダム値を設定する. このように離散化した ABC を DBN の学習方法として用いる.

6.2 BN の構造表現

本研究では簡単のためトランジションネットワークの構造のみを学習対象とする. 時刻 T , $T+1$ におけるトランジションネットワークがあるとき, この構造に対して列を子ノード($T+1$), 行を親ノード(T) とする関係表を作成し子ノードと親ノードのリンクの有無をバイナリ値で表す. そして前述の評価基準を用いて評価値の高いビット列の探索を行う.

7. シミュレーション実験

離散化した ABC での BN の構造学習の有効性を検証するため, ベンチマークとして以下に示す学習データに対して実験を行った. また評価基準は BIC を用いた.

7.1 実験環境

図 2 に学習対象の構造を示す. 時刻 T の初期値はすべてランダムなバイナリ値をとるように決定した. 時刻 T の親ノードが持つ値を足した値が 0 以上なら $T+1$ の子ノードは 1 になり, 0 の時はランダムでバイナリ値を取るような BN の構造を学習する.

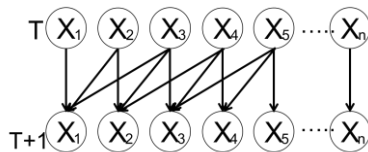


図 2 実験ネットワーク構造

7.2 実験 1

提案手法の有効性を検証するため DPSO, GA, および提案手法の性能を比較する. 10 変数(次元数: $10 \times 10=100$), 学習データ数 300, イタレーション回数を 1000, 試行回数を 10 とした. 表 1 に各手法のパラメータを示す.

パラメータ	DPSO	GA	提案手法
個体数	400	400	400
ρ	n/a	n/a	0.4
Limit	n/a	n/a	20000
w	0.5	n/a	n/a
c1, c2	2	n/a	n/a
交叉方法	n/a	二点交叉	n/a
突然変異率	n/a	0.01	n/a

表 1 実験 1 パラメータ

GA, DPSO のパラメータ, 提案手法の Limit 値は事前実験から決定した. 図 3 に実験結果を示す. この結果から提案手法が局所解に陥らず DPSO, GA よりも早く良解を求められていることが分かる.

7.3 実験 2

8 変数(次元数: $8 \times 8=64$), データ数 200, 個体数 400 において提案手法における式(3)の ρ の値を 0, 0.25, 0.5, 0.75, 1 に変更した時の実験を行う. limit 値は 12800, イタレーション回数を 800, 試行回数を 10 回とした. 図 4 に実験結果を

示す. この結果から ρ が 1 のときに, より良い学習ができていることがわかる. これは ABC の探索では 1 次元のみ変化させるので現探索点とランダムで選んだ探索点のある次元が同じ値であるときは反転させることで確実に現探索点の構造を変化させることができるためだと考えられる.

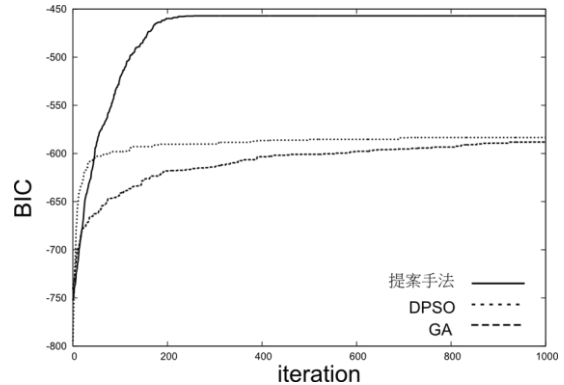


図 3 DPSO, GA との比較

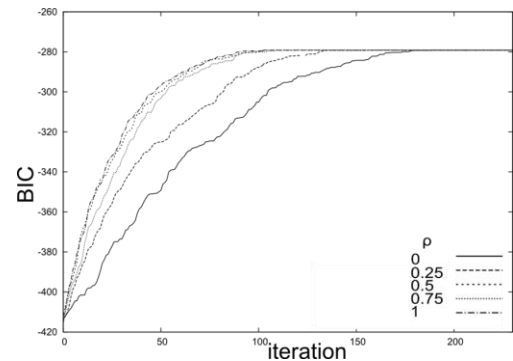


図 4 ρ を変化させた時のグラフ

8. おわりに

本研究では人工蜂コロニー最適化を用いた DBN の構造学習方法を提案し, 各種法との比較を行いその有効性を検証した. また式(3)の ρ の値を変化させた実験を行い, 学習性能に与える影響について検証した.

参考文献

- [1] J. Pearl "Probabilistic Reasoning in Intelligent System", Morgan Kaufmann (1988).
- [2] K. Murphy "Dynamic Bayesian Networks: Representation, Inference and Learning", PhD thesis, Computer Science Division UC Berkeley, 2002.
- [3] R. E. Steuer, Multiple Criteria Optimization: Theory, Computations, and Application. New York: John Wiley & Sons, Inc., ISBN 047188846X, 1986 BIC
- [4] P. Larranga, M. Poza, "Structure Learning of Bayesian Networks by Genetic Algorithms: A Performance Analysis of Control parameters", pattern analysis and machine intelligence, vol. 18 pp. 912-926, Sep 1996
- [5] W. Guo, X. Gao, Q. Xiao, "Bayesian Optimization Algorithm for Learning Structure of Dynamic Bayesian Networks from Incomplete Data", Proc. CCDC 2008, pp. 2088-2093, 2008
- [6] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm", Applied Soft Computing, vol. 8, pp. 687-697, 2008.