

Bayesian Network の理論調査
中村 康昭

1 はじめに

これまで, Goldberg らのレポート¹⁾を参考にして BOA の実装を試みてきたが, レポートに記されているものと同等の性能を得ることができなかったため, 他の文献を用いて Bayesian Network の確率の伝搬アルゴリズム, ネットワークの構成アルゴリズムの理論について, 調査を行った. 本稿ではその内容について述べる.

2 進捗状況

2.1 確率の伝搬

ベイジアンネットワークでは, 各ノードが変数を示し, その依存関係が矢印によって示される. 本節では, ネットワークのトポロジーを単一結合とし, 確率伝搬についての説明を行う.

Fig. 1 に, 構築されたと仮定するネットワークを示す.

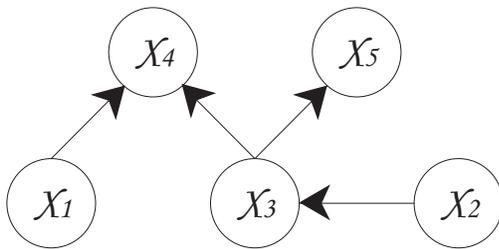


Fig. 1 Bayesian Network の例

このとき, 各ノードの条件付き確率表 (CPT) は求まっているものとする. $x_2 = e_2, x_5 = e_5$ という値が観測されたときに, 残ノードの値の発現確率がどのように変化するかを求めるのが確率伝搬となる.

ここで, 注目するノードを z と置き, どのように確率が伝搬するかを見る. ノード z の上流 (親ノード群) にて得られた証拠を e^+ とし, 下流 (子ノード群) にて得られた証拠を e^- と表記する.

確率の伝搬は式 (1) に示される式によって定義される.

$$\begin{aligned}
 P(X_j|e) &= P(z|e^+, e^-) \\
 &= \frac{P(e^-|z, e^+)P(z|e^+)}{P(e^-|e^+)} \quad (1)
 \end{aligned}$$

Fig. 1 における例では証拠として与えられるのが x_2 と x_5 のみであり, これ以外のノードについて確率を伝搬させる. この例では入れ子構造になっているととらえ,

式 (1) を繰り返し使うことによって確率が伝搬していくことになる.

2.2 ネットワーク構築

現在よく知られているベイジアンネットワークの構造学習アルゴリズムとして, K-2 アルゴリズムがある. 以前の報告で K2 Metric の報告を行ったが, 本アルゴリズムとは関係のないものであるため, 注意する必要がある.

全ての結合可能なノードについて, その評価を行うために結合を行ってはいは組み合わせ爆発を起こすため, K-2 アルゴリズムを用いて実用的な時間でネットワークの評価を行う. 具体的な手順を以下に示す.

1. 各ノードについて, 親になるノードを限定する
2. 候補となる親ノードを一つずつ加えてネットワークを作る
3. 評価が高くなったときには親ノードとして採用する
4. 候補がなくなるか, 加えても評価が高くならなくなるまで繰り返す
5. 全てのノードについて以上を繰り返す

生成したグラフ構造を評価する際には, 過剰な複雑性を避けるため, データの尤度と構造の複雑さの両面を考慮する必要が生じる. モデルを選択するためには MDL 基準や, そのほか様々なものが採用されている. K2 Metric とは, ここで用いられる評価基準である.

3 今後の予定

調査より, これまでのベイジアンネットワーク構築方法が BOA のものと異なるという証拠が得られなかった. 現在プログラムを修正中であり, これを完成させ, 再度検証を行う予定である.

参考文献

- 1) David E. Goldberg Martin Pelikan and Erick Cantu-Paz. BOA: The bayesian optimization algorithm. Technical Report 99003, IlliGAL Report, 1999.
- 2) 鈴木 譲, ベイジアンネットワークにおける確率伝搬について, 2001 年ベイジアンネットチュートリアル
- 3) 木村陽一, 不確実性モデリングのための情報表現: ベイジアンネット, 2001 年ベイジアンネットチュートリアル