

ベイジアンネットモデル構築の実際と

BAYONET 次期バージョンのご紹介

(株)数理システム 科学技術部

石富 妙

1. はじめに

ベイジアンネットワークは様々な事象間の因果関係(厳密には確率的な依存関係)をグラフ構造で表現するモデリング手法の一つで、故障診断、気象予測、医療的意思決定支援、マーケティング、推薦システムなど様々な分野で利用や研究が行われています。

ベイジアンネットワークに興味があり実際にモデルを作りたいけど、どうしたらよいかわからない、という声をよく聞きます。今回の発表では、ベイジアンネットモデル構築のプロセスを BAYONET のデモを交えて解説します。併せて BAYONET の次期バージョンについてもご紹介します。

2. ベイジアンネットモデルについて

ベイジアンネットワークはベイズの定理を基本としています。ベイズの定理は二つの事象について直接の依存関係を表していますが、これをネットワークへと拡張したのがベイジアンネットワークです。ベイジアンネットワークでは事象をノードで表します。事象間に直接の確率的な依存関係があれば対応するノードを矢印で結びます。またその依存関係は条件付き確率表で定量的に表現されます。ネットワーク構造は非循環有向グラフ(DAG)でなければなりません。

ベイジアンネットモデルは次のような特徴があります。

- ネットワークで表現されるため構造の理解が容易
ベイジアンネットモデルはグラフィカルモデルです。ノード間の依存関係を視覚的にわかりやすく表現できます。
- 予測を行うとき、説明変数の入力に欠損があってもよい
ニューラルネットワークなどと異なり、ベイジアンネットモデルは予測時に全ての説明変数に観測値を入力する必要はありません。観測値のない変数については条件に応じた適切な確率分布が割り当てられます。
- モデルの利用用途が限定されない

ベイジアンネットワークモデルは矢印の順方向だけでなく、逆方向にも推論が行えます。よって観測を入力する変数と予測対象となる変数がモデルで限定されることなく、自由に選択できます。

3. ベイジアンネットワークモデル構築の実際

ベイジアンネットワークモデルは次のような工程を経て実用化されます。ただし各工程で逐次フィードバックが行われるため、実際にはスパイラルな過程になります。学習データからモデルを作成する場合には[データの前処理]が必要となる場合があります。

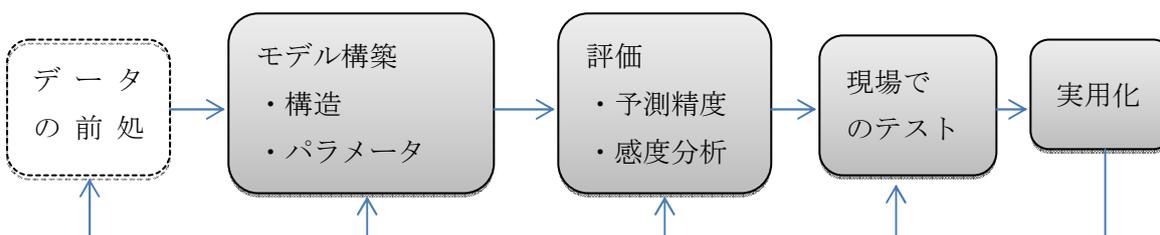


図 1 モデルの実用化までのプロセス

以下で[データの前処理]から[評価]までについて説明します。

- データの前処理

モデルに使用する変数を精査します。BAYONET は数値データを直接扱えないため、適切な分割手法を用いて離散化を行う必要があります。(離散化は BAYONET の[モデルの自動構築]で自動的行われます。)

- モデル構築

この過程ではネットワーク構造と条件付き確率表を決定します。学習データがあれば構造探索アルゴリズムにより構造が得られます。また条件付き確率表もデータから適切な分布を得ることができます。

ベイジアンネットワークの構造探索アルゴリズムには、評価基準により変数間の依存関係を抽出するスコアリングベースの手法と、変数間の独立性判定により構造を定める CI ベースの手法があります。BAYONET の構造探索アルゴリズムはスコアリングベースの手法を採用しています。

さらにエキスパートの知識があれば、構造探索の制約条件として活用することができます。構造探索アルゴリズムは、変数間の確率的な依存関係を抽出するものですので、因果関係とは必ずしも一致しません。例えば因果の向きとは逆向きにリンクが引かれる場合もあります。エキスパートの知識をモデル構築の制約条件として与えることで、因果の方向が正しいより自然なモデルを構築することができます。

- 評価

ベイジアンネットワークモデルの評価手法には[予測精度]及び[感度分析]の二つがあります。ここでは BAYONET に実装されている[予測精度]について説明します。

- ・ 予測精度(正解率)

検証用のデータからモデルに観測値を入力して推論を行います。目的とする変数について予測値と観測値(検証データの値)の一致する割合を算出します。

4. BAYONET の機能と次期バージョンのご紹介

BAYONET は(独)産業技術総合研究所によって開発され、数理システムが販売を行っているベイジアンネットワークモデルのためのソフトウェアです。BAYONET は以下のような機能が実装されています。

- ベイジアンネットワークモデルの自動構築

Greedy Search アルゴリズムによりモデルの構造探索を行います。探索はマルチスレッドに対応しており、マルチコア環境を活かすことができます。

探索過程では双方向リンクや循環が生成されますが、自動的に除去します。

- 確率推論機能

loopyBP、MSSM 法により確率推論を行います。またエクセルのアドインを使って、複数のエビデンスデータに対してまとめて推論を実行することも可能です。

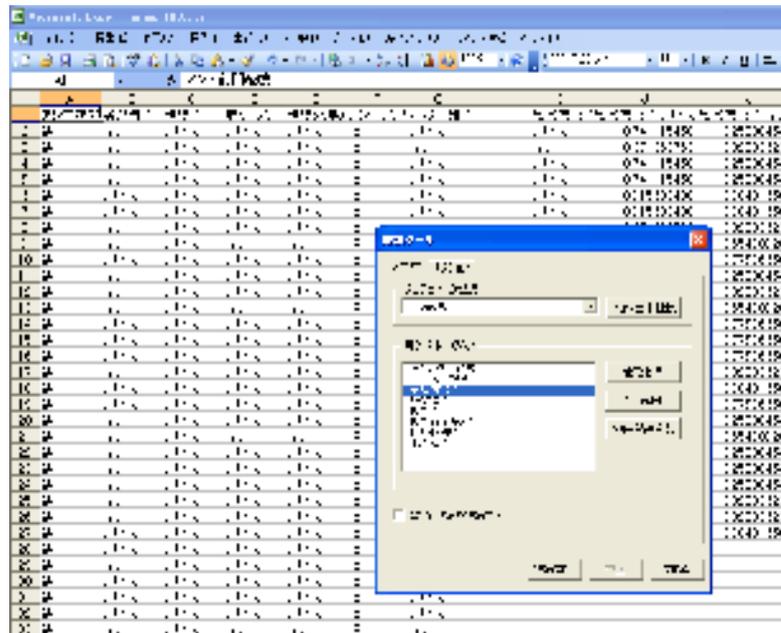


図 2 エクセルアドインによる推論

- モデル検証機能
ベイジアンネットワークモデルの予測精度や平均対数尤度などを算出します。



図 3 モデルの検証ツール

- 64 ビットに対応
バージョン 5.0.2 から 64 ビット環境にも対応しており、より規模の大きなモデルの構築が可能になりました。

BAYONET の次期バージョンは主に GUI について変更を行います。

- 情報パネル
画面左側の情報パネルを一新します。
- モデルの自動構築
構造探索のウィザードを完全に日本語に対応した新しいウィザードへ変更します。
モデル構築の設定が保存できるなど、新しい機能も搭載します。