

漸近一致性をもつベイジアンネットワーク IRT の  
制約ベース構造学習

2022年2月22日

情報数理工学プログラム

学籍番号 1810723

青木 健登

指導教員 植野 真臣

# 1 まえがき

近年、TOEIC や情報処理技術者試験など、様々な大規模テストが e テスティングにより実施されている。e テスティングは、項目反応理論 (Item Response Theory : IRT)[1] とよばれるテスト理論に基づき、項目データベースを構築する。IRT は、被験者が項目に正答する確率を、被験者の能力を表す潜在変数の関数として表現した数理モデルであり、(1) テストの出題内容に依らず被験者を同一尺度上で評価できる、(2) 既存の項目から構成されたテストを実施する前に得点分布を予測できる、等の利点を有する。

IRT モデルは、能力潜在変数を所与として各項目への正答確率が条件付き独立になる、という仮定を前提としている。この仮定を「局所独立性」という。一方で、実際のテストでは局所独立性の仮定が成り立たないことが多く [2]、これを局所依存という。通常の IRT モデルを局所依存性をもつ項目群に対して適用すると、項目パラメータの推定値の偏り [3] や、能力推定値の標準誤差の増加 [4] などの問題を引き起こすことが知られている。

上記の問題を解決するために、Ueno[5] は従来の IRT モデルにおける局所独立性の仮定を緩和したベイジアンネットワーク IRT を提案している。ベイジアンネットワーク IRT では、被験者の能力パラメータを連続潜在変数とするベイジアンネットワーク構造で項目間の局所従属性を表現する。ベイジアンネットワーク IRT は、従来の 2 パラメータ・ロジスティックモデル (2-parameter logistic model : 2PL モデル) に比べて、従属性のある未知の項目に対する被験者の反応の予測性能が高い [5]。したがって、ベイジアンネットワーク IRT の構造学習を行うことで、項目間の局所従属性を同定し、従属性が未知の項目群に対する被験者の反応の予測精度を向上できる。

ベイジアンネットワーク IRT の学習法として、統計的検定を用いて、橋本ら [6] により項目潜在構造分析 (Item Latent Structure Analysis : ILS 分析) が提案されている。ILS 分析は、潜在条件付き独立性検定 (Latent Conditional Independence Test : LCI 検定) と呼ばれる条件付き独立性検定 (Conditional Independence Test : CI テスト) [7] を用いて学習を行う。LCI 検定の精度は設定する閾値に強く影響を受けるため、これを用いた学習は漸近一致性をもたない。そのため、ILS 分析によるベイジアンネットワーク IRT の構造学習は、漸近的に真の構造を学習できる保証がないという問題がある。

ベイジアンネットワークの分野では、Natori ら [8] は漸近一致性をもつ構造学習アルゴリズムとして、ベイズファクター (Bayes Factor : BF) を用いる検定を組み込んだ制約ベースの構造学習アルゴリズムを提案している。本論文では、Natori ら [7] のアイデアをベイジアンネットワーク IRT に適用し、能力潜在変数について周辺化した BIC (Bayesian Information Criterion) による BF を提案する。能力潜在変数の推定はニュートン法や MCMC 法などの繰り返し計算を必要とし、計算コストが高い。そこで、提案手法では能力潜在変数について周辺化する。この手法でも、項目パラメータを繰り返し計算により求める必要があるため、1 回あたりの CI テストに要する計算時間はまだ大きく、大規模ネットワークを学習できないという問題がある。

ここで、ベイジアンネットワーク IRT の性質として、真のベイジアンネットワーク IRT はベイ

ジアンネットワークの厳密学習で得られるグラフの部分グラフになることが証明される。この性質を利用し、本論文では、最初にベイジアンネットワークの厳密学習を行い、その結果削除されなかったエッジに対してのみ、提案の BF を用いた CI テストを行う。この方法により、完全グラフから制約ベース構造学習を開始する場合に比べ、提案の BF を用いた CI テストの回数を削減できる。

以上より、本論文の提案手法は、次の 2 つの学習から構成される。

1. ベイジアンネットワークの厳密学習を行う。
2. 真のベイジアンネットワーク IRT が 1 で学習されたグラフの部分グラフであることを制約として、提案の BF を用いた CI テストによるベイジアンネットワーク IRT の制約ベースの構造学習を行う。

提案手法の利点として、以下の 2 つが挙げられる。

1. 構造学習の計算コストを削減できる。
2. 漸近一致性をもつため、データ数が十分に大きいとき、真の構造を学習できる。

シミュレーション実験により、ベイジアンネットワークの厳密学習で得られるグラフの部分グラフ制約を用いてベイジアンネットワーク IRT を学習することで、制約を用いない場合に比べ構造学習全体の計算コストを削減できることを示す。また、データ数が十分に大きいとき、提案手法が従来手法よりも学習の精度が高いことを示す。

## 2 項目反応理論

項目反応理論 (Item Response Theory : IRT) [1] は、項目に被験者が正答する確率を被験者の能力を表す潜在変数  $\theta$  の関数として表現する数理モデルである。通常の二値反応形式の IRT では、被験者  $i$  のテストにおける  $j$  番目の項目への反応は、正答なら確率変数  $u_{ij}$  が 1、誤答なら  $u_{ij}$  が 0 となることで表現する。例えば、最も広く使われている IRT モデルである 2 パラメータ・ロジスティックモデル (2-parameter logistic model : 2PL モデル) では、能力潜在変数が  $\theta_i$  である被験者  $i$  が項目  $j$  に正答する確率は次のように表される。

$$P(u_{ij} = 1 | \theta_i, a_j, b_j) = \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_j))}$$

- $a_j$  : 項目  $j$  の識別力パラメータ
- $b_j$  : 項目  $j$  の難易度パラメータ

識別力パラメータ  $a_j$  は、 $\theta_i = b_j$  付近の被験者の能力をどれほどよく識別できるかを表すパラメータである。 $a_j$  の値が大きいほど、 $\theta_i = b_j$  付近の能力値の変化により、正答確率の変化が大きくなる。難易度パラメータ  $b_j$  は項目の難易度を表すパラメータである。 $b_j$  の値が大きいほど、被験者が項目  $j$  に正答する確率は低くなる。識別力パラメータと難易度パラメータをまとめて項目パラメータという。

IRT は、(1) テストの出題内容に依らず被験者を同一尺度上で評価できる、(2) 既存の項目から構成されたテストを実施する前に得点分布を予測できる、等の利点を有する。

## 2.1 局所独立性の仮定

2PL モデルの尤度関数は次のように定義される。

$$L(U|\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) = \prod_{i=1}^I \prod_{j=1}^J \left[ \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_j))} \right]^{u_{ij}} \cdot \left[ 1 - \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_j))} \right]^{1-u_{ij}} \quad (1)$$

- $U = \{u_{ij}\}$  ( $i = 1, \dots, I; j = 1, \dots, J$ )
- $\boldsymbol{\theta} = \{\theta_i\}$  ( $i = 1, \dots, I$ )
- $\mathbf{a} = \{a_j\}$  ( $j = 1, \dots, J$ )
- $\mathbf{b} = \{b_j\}$  ( $j = 1, \dots, J$ )
- $I$ : 被験者の総数
- $J$ : 項目の総数

式 (1) の右辺は、各項目に対する条件付き確率の積を、すべての被験者について掛け合わせたものである。この式は、異なる項目に対する被験者の反応が  $\theta$  を所与として条件付き独立であるという仮定により導かれる。この仮定は項目の「局所独立性」と呼ばれる。しかし、実際のテストでは、様々な要因により局所独立性の仮定を満たさない項目群が含まれることがある [2]。例えば、次のような場合が挙げられる [4]。

- 2 項目間に直接的な繋がりが存在する  
(例 前問に対する反応を使って解くような問題)
- 2 項目間に間接的な繋がりが存在する  
(例 ある長文に関連する複数の問題)
- 正答確率に能力潜在変数  $\theta$  以外の特性が反映される  
(例 テスト分析者が仮定した  $\theta$  以外の特性 (性別, 国籍) によって正答確率が変化する場合)

このような局所独立性の仮定を満たさないデータに対して通常の IRT モデルを適用すると、識別力パラメータの過大推定 [3] や能力パラメータの推定値の標準誤差 [4] が大きくなるなどの問題が知られている。

## 2.2 ベイジアンネットワーク IRT

Ueno[5] は、前節の偏ったパラメータの値を推定してしまう問題を解決するために、IRT の局所独立性の仮定を緩和したベイジアンネットワーク IRT を提案している。ベイジアンネットワーク IRT では、項目間の局所従属性を被験者の能力パラメータを連続潜在変数とするベイジアンネットワーク構造で表現する。このモデルでは、従属性がある項目に対する反応によって、他の項目

の難易度パラメータが変化する。

例えば、図1のようなグラフで表される3項目のテストを考える。この項目群のうち、項目1は独立、項目2と3は従属である。ここで、他の項目の難易度に影響を与える項目を親項目、難易度が変わる項目を子項目と呼ぶ。この構造では、項目2が項目3の親項目、項目3が項目2の子項目となる。ベイジアンネットワークIRTでは、子項目は難易度パラメータを複数保持しており、親項目に対する反応パターンによって難易度が変わる。図1の例では、親項目である項目2に正答すると項目3の難易度が下がり、項目2に誤答すると項目3の難易度が上がる。すなわち、項目2に正答する人は項目3に正答しやすく、項目2に誤答する人は項目3に誤答しやすいことをモデル化している。

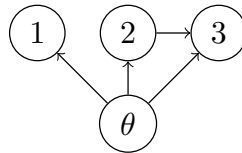


図1 ベイジアンネットワークIRTの例

ベイジアンネットワークIRTでは、ベイジアンネットワーク構造を  $G$  とすると、被験者  $i$  の全項目への反応パターンの条件付き同時確率は以下の式で表される。

$$P(\mathbf{u}_i | \theta_i, \mathbf{a}, \mathbf{b}, G) = \prod_{j=1}^J \prod_{k=1}^{2^{m_j}} \left( \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_{jk}))} \right)^{u_{ij} \cdot u_{ijk}} \cdot \left( 1 - \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_{jk}))} \right)^{1 - u_{ij} \cdot u_{ijk}}$$

- $\mathbf{u}_i = \{u_{ij}\} (j = 1, \dots, J)$
- $\mathbf{a} = \{a_j\} (j = 1, \dots, J)$
- $\mathbf{b} = \{b_{jk}\} (j = 1, \dots, J; k = 1, \dots, 2^{m_j})$
- $m_j$  : 項目  $j$  の親項目の数

ここで、 $u_{ijk}$  は被験者  $i$  の項目  $j$  の親項目群に対する反応パターンが  $k$  番目のとき1、そうでないとき0となる変数であり、 $b_{jk}$  は被験者が項目  $j$  の親項目群に対して  $k$  番目の反応パターンを示したときの項目  $j$  の難易度パラメータである。

ベイジアンネットワークIRTでは、項目  $j$  のパラメータ数は  $2^{m_j} + 1$  個となる。つまり、項目  $j$  の親項目数が増えるにしたがって項目  $j$  の項目パラメータ数は指数的に増加する。

ベイジアンネットワークIRTは、従来の2PLモデルに比べて、従属性のある未知の項目に対する被験者の反応の予測性能が高いことが知られている [5]。

したがって、ベイジアンネットワークIRTの構造学習を行うことで、項目間の局所従属性をベイジアンネットワーク構造として同定し、従属性が未知の項目群に対する被験者の反応の予測精度を向上できる。

### 3 ベイジアンネットワーク IRT の構造学習

橋本ら [6] は、ベイジアンネットワーク IRT の構造学習法として、項目潜在構造分析 (Item Latent Structure Analysis : ILS 分析) を提案している。ILS 分析では、潜在変数を所与とする条件付き独立性検定 (Conditional Independence test : CI テスト) である潜在条件付き独立性検定 (Latent Conditional Independence Test : LCI 検定) [7] を用いて学習を行う。

LCI 検定で 2 変数  $X_j$  と  $X_{j'}$  の条件付き独立性を検定する方法を説明する。まず、完全有向グラフ  $G_c$  を考える。エッジの向きは、テストで出題される順番が前の項目から後の項目への向きを仮定する。 $G_c$  における  $X_j$  と  $X_{j'}$  の親変数集合を  $\mathbf{X}^{-jj'}$  とする。次に、 $\mathbf{X}^{-jj'}$  を所与とする  $X_j$  と  $X_{j'}$  の条件付き相互情報量 (Conditional Mutual Information : CMI) である  $I(X_j, X_{j'} | \mathbf{X}^{-jj'}, G_c)$  を求める。実際には、その推定量として次の LCI 指標を計算する。

$$I_{jj'} = \frac{1}{N} \sum_i \sum_{x_j=0}^1 \sum_{x_{j'}=0}^1 N_{x_j x_{j'} i} \log_2 \frac{N_{x_j x_{j'} i} N_i}{N_{x_j i} N_{x_{j'} i}} \quad (2)$$

ここで、 $N$  は被験者数、 $N_i$  は  $\mathbf{X}^{-jj'} = \mathbf{x}_i^{-jj'}$  という反応パターンを示した被験者の人数、 $N_{x_j, i}$  は  $X_j = x_j$  かつ  $\mathbf{X}^{-jj'} = \mathbf{x}_i^{-jj'}$  という反応パターンをした被験者の人数、 $N_{x_j x_{j'} i}$  は  $X_j = x_j$ ,  $X_{j'} = x_{j'}$  かつ  $\mathbf{X}^{-jj'} = \mathbf{x}_i^{-jj'}$  という反応パターンを示した被験者の人数である。LCI 検定では、LCI 指標が閾値  $\varepsilon$  以上のとき  $X_j$  と  $X_{j'}$  は能力潜在変数  $\theta$  を所与として条件付き従属、そうでないとき条件付き独立と判定する。Hashimoto ら [7] は、閾値として 0.01, 0.05, 0.1 を用いている。

ILS 分析は、LCI 検定をすべての 2 変数の組  $(X_j, X_{j'})$  ( $j = 1, \dots, j' - 1, j' = 2, \dots, J$ ) に対して適用することで、ベイジアンネットワーク IRT の構造学習を行う。ILS 分析のアルゴリズムを Algorithm1 に示す。

しかし、LCI 検定の精度は設定する閾値に強く影響を受けるため、漸近一致性をもたないという問題がある。

## 4 提案手法

本論文では、計算コストを考慮した上で、漸近一致性をもつベイジアンネットワーク IRT の構造学習アルゴリズムを提案する。提案手法は、[第 1 段階目の学習] ベイジアンネットワークの厳密学習、[第 2 段階目の学習] BIC (Bayesian Information Criterion) による BF を用いた学習、の 2 つの学習から構成される。

### 4.1 BIC による BF

ベイジアンネットワークの分野では、Natori ら [8] は Steck ら [9] が提案した BF を用いる CI テストを制約ベースアルゴリズムに組み込み、漸近一致性をもつ構造学習を実現している。本節では、BF を用いる CI テストをベイジアンネットワーク IRT に適用し、BIC による BF を用いる CI テストを提案する。

---

**Algorithm 1** ILS 分析

---

**Require:**  $N$  ノード変数, データ  $U$ , 閾値  $\varepsilon$

**Ensure:** ベイジアンネットワーク構造  $G$

```
1:  $G_c \leftarrow$  完全有向グラフ
2: for 各ノード  $X_{j'}$  do
3:   for  $G_c$  における  $X_{j'}$  の各親ノード  $X_j$  do
4:      $I_{jj'} \leftarrow$  項目  $j$  と  $j'$  の LCI 指標
5:   end for
6: end for
7:  $G \leftarrow G_c$ 
8: for 各ノード  $X_{j'}$  do
9:   for  $X_{j'}$  の各親ノード  $X_j$  do
10:    if  $I_{jj'} < \varepsilon$  then
11:       $G \leftarrow G$  からエッジ  $(X_j, X_{j'})$  を削除したグラフ
12:    end if
13:   end for
14: end for
15: return  $G$ 
```

---

BF は、2 つのモデルの周辺尤度の比で表され、ベイジアンネットワークにおいて厳密なモデル選択が可能となる指標である。Steck ら [9] は、2 ノード間が独立・従属なモデルの周辺尤度の比による BF を用いた CI テストを提案した。例として、 $X$  と  $Y$  間について各ノードの共通の親ノード集合を  $\mathbf{Z}$  としたときの従属なモデルを  $g_1$ 、独立なモデルを  $g_2$  とする。このときの BF を  $\text{BF}(X, Y | \mathbf{Z})$  とし、その対数である対数 BF を  $\log \text{BF}(X, Y | \mathbf{Z})$  とすると

$$\log \text{BF}(X, Y | \mathbf{Z}) = \log P(\mathbf{D} | g_1, \alpha) - \log P(\mathbf{D} | g_2, \alpha) \quad (3)$$

と表される。ここで、 $P(\mathbf{D} | g, \alpha)$  はデータを  $\mathbf{D}$ 、ハイパーパラメータを  $\alpha$  とした時のベイジアンネットワーク構造  $g$  の周辺尤度である。Steck らでは、 $P(\mathbf{D} | g, \alpha)$  に Bayesian Dirichlet equivalence uniform (BDeu) と呼ばれるスコアを用いる。BDeu におけるハイパーパラメータ  $\alpha$  を Equivalent Sample Size (ESS) という。Steck ら [9] は、式 (3) が 0 以上か否かで  $g_1$ 、 $g_2$  のどちらの構造を選択するか判定する CI テストを提案した。

Natori ら [8] は、Steck らの提案した BF を用いる CI テストを RAI アルゴリズム [10] に組み込んだベイジアンネットワークの制約ベース構造学習を提案している。Natori らの手法は漸近一致性を有し、大規模ネットワークの構造学習を可能にする。

したがって、ベイジアンネットワーク IRT の制約ベース構造学習においても、BF を CI テストに用いることで漸近一致性をもつ構造学習が期待できる。しかし、離散変数のみを扱う通常のベイジアンネットワークと異なり、ベイジアンネットワーク IRT には連続変数  $\theta$  が含まれる。そのため、Natori ら [8] が用いている BDeu による BF をそのままベイジアンネットワーク IRT に適

用することは困難である。

ベイジアンネットワーク IRT に適用するための BF を考える。識別力パラメータ  $\mathbf{a}$ 、難易度パラメータ  $\mathbf{b}$  をまとめて  $\boldsymbol{\xi}$  と書くと、ベイジアンネットワーク IRT におけるベイジアンネットワーク構造  $G$  の周辺尤度  $p(U|G)$  は

$$p(U|G) = \iint g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G) d\boldsymbol{\xi} d\boldsymbol{\theta} \quad (4)$$

と表される。ここで  $g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G)$  は次のように表される。

$$g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G) = L(U|G, \boldsymbol{\xi}, \boldsymbol{\theta}) f(\boldsymbol{\xi}) f(\boldsymbol{\theta})$$

$f(\boldsymbol{\xi})$ ,  $f(\boldsymbol{\theta})$  はそれぞれ項目パラメータ  $\boldsymbol{\xi}$ 、能力パラメータ  $\boldsymbol{\theta}$  の事前分布である。 $\mathbf{a}$  と  $\mathbf{b}$  は i.i.d. を仮定しているため、 $\mathbf{a}$ ,  $\mathbf{b}$  の事前分布をそれぞれ  $f(\mathbf{a})$ ,  $f(\mathbf{b})$  とすると  $f(\boldsymbol{\xi}) = f(\mathbf{a})f(\mathbf{b})$  である。周辺尤度  $p(U|G)$  を用いて、ベイジアンネットワーク構造  $G_1$ ,  $G_2$  の対数 BF は次のように表される。

$$\log \text{BF}(G_1, G_2) = \log p(U|G_1) - \log p(U|G_2)$$

しかし、式 (4) の周辺尤度は解析的に解くことができない。一方、モンテカルロ法などの数値積分を適用する場合、周辺尤度はパラメータ  $\mathbf{a}, \mathbf{b}, \boldsymbol{\theta}$  についての重積分で表されるため計算量が膨大になる。

そこで提案手法では、式 (4) の周辺尤度の対数を近似する。まず

$$\boldsymbol{\xi}_{\text{MAP}} = \arg \max_{\boldsymbol{\xi}} g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G)$$

とおくと、 $g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G)$  の  $\boldsymbol{\xi}$  に関する積分は次のようにラプラス近似できる。

$$\int g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G) d\boldsymbol{\xi} \simeq g(U, \boldsymbol{\xi}_{\text{MAP}}, \boldsymbol{\theta}|G) (2\pi)^{\frac{k}{2}} \det(-H_{\boldsymbol{\xi}_{\text{MAP}}})^{-\frac{1}{2}}$$

ここで、 $H_{\boldsymbol{\xi}_{\text{MAP}}}$  は  $\log g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G)$  の  $\boldsymbol{\xi}_{\text{MAP}}$  におけるヘッセ行列、 $k$  は項目パラメータの数で  $k = J + \sum_{j=1}^J 2^{m_j}$  である。両辺を  $\boldsymbol{\theta}$  で積分し、対数をとると

$$\begin{aligned} & \log \iint g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G) d\boldsymbol{\xi} d\boldsymbol{\theta} \\ & \simeq \log \int L(U|G, \boldsymbol{\xi}_{\text{MAP}}, \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} + \log f(\boldsymbol{\xi}_{\text{MAP}}) + \frac{k}{2} \log(2\pi) - \frac{1}{2} \log \det(-H_{\boldsymbol{\xi}_{\text{MAP}}}) \end{aligned} \quad (5)$$

を得る。式 (5) は、サンプル数  $N$  が十分大きいとき

$$\log \iint g(U, \boldsymbol{\xi}, \boldsymbol{\theta}|G) d\boldsymbol{\xi} d\boldsymbol{\theta} \simeq \log \int L(U|G, \boldsymbol{\xi}_{\text{MAP}}, \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} - \frac{k}{2} \log N$$

と近似できる。この右辺の  $\boldsymbol{\xi}_{\text{MAP}}$  を MAP 推定値  $\hat{\boldsymbol{\xi}}_{\text{MAP}}$  に置き換えたものを BIC といい

$$\text{BIC}(G) = \log \int L(U|G, \hat{\boldsymbol{\xi}}_{\text{MAP}}, \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} - \frac{k}{2} \log N \quad (6)$$



---

**Algorithm 2** 提案の BF を用いる条件付き独立性検定関数

---

**Require:** 検定するエッジの子ノード  $X_i$ , 検定するエッジの親ノード  $X_j$ , 所与とするノード集合  $c$ , 仮定するベイジアンネットワーク構造  $G$ , データ  $U$

**Ensure:**  $\theta$  と  $c$  を所与として  $X_i, X_j$  が条件付き独立なら 1, 条件付き独立でないなら 0

- 1:  $G_{nd} \leftarrow G$  における  $X_i$  の親ノード集合を  $\{X_j\} \cup c$  に変更したグラフ
  - 2:  $BIC_{nd} \leftarrow BIC(G_{nd}, U)$  : ただし,  $BIC(G, U)$  はデータ  $U$  に対するグラフ  $G$  の BIC を返す関数.
  - 3:  $G_d \leftarrow G_{nd}$  からエッジ  $(X_i, X_j)$  を削除したグラフ
  - 4:  $BIC_d \leftarrow BIC(G_d, U)$
  - 5: **if**  $BIC_d - BIC_{nd} > 0$  **then**
  - 6:     **return** 1
  - 7: **end if**
  - 8: **return** 0
- 

と表す.  $\theta$  の推定はニュートン法や MCMC 法などの繰り返し計算を必要とし, 計算コストが高い. そこで, 提案手法では能力潜在変数について周辺化する.  $\xi_{\text{MAP}}$  は,  $\theta$  について周辺化した事後分布に対して EM アルゴリズムを適用して求める.

BIC を用いて, 対数 BF は次のように近似できる.

$$\log \text{BF}(G_1, G_2) \simeq \text{BIC}(G_1) - \text{BIC}(G_2) \quad (7)$$

提案手法では, 式 (7) で表される BIC による BF を検定統計量として用いる.

提案の BF を用いて項目の集合  $Z$  と  $\theta$  を所与とした 2 項目  $X_i$  と  $X_j$  の条件付き独立性を検定する方法を説明する. その時点で仮定しているベイジアンネットワーク構造を  $G$  とする. Hashimoto ら [7] の LCI 検定と同様に, エッジの向きはテストで出題される順番が前の項目から後の項目への向きを仮定する. 次の手順により条件付き独立性を検定する.

1.  $G$  において  $X_i$  の親項目集合を  $Z$  に置き換えたグラフを  $G_1$  とする
2.  $G$  において  $X_i$  の親項目集合を  $Z \cup X_j$  に置き換えたグラフを  $G_2$  とする
3.  $\log \text{BF}(G_1, G_2) > 0$  ならば  $\theta$  と  $Z$  を所与として  $X_i$  と  $X_j$  は条件付き独立, そうでなければ  $\theta$  と  $Z$  を所与として  $X_i$  と  $X_j$  は条件付き独立でない

式 (6) の第一項は  $\theta$  について周辺化した尤度の対数であるため, この BIC スコアは分解可能でない. すなわち, 検定においてベイジアンネットワーク構造  $G_1$  全体の BIC とベイジアンネットワーク構造  $G_2$  全体の BIC の差を求める必要がある. 提案の BF を用いる CI テストを行う関数を Algorithm2 に示す.

本論文では, 提案の BF を用いる CI テストを一般的な制約ベースアルゴリズム [11],[12] に組み込む. これを第 2 段階目の学習と呼ぶ.

## 4.2 CI テスト回数の削減のためのベイジアンネットワークの学習

前節で提案した BIC は、項目パラメータを繰り返し計算により求める必要があるため、1 回あたりの CI テストに要する計算時間はまだ大きく、大規模ネットワークを学習できないという問題がある。この問題を解決するために、提案手法では、真のベイジアンネットワーク IRT がベイジアンネットワークの厳密学習結果のグラフの部分グラフになる制約を用いることで、ベイジアンネットワーク IRT の解の探索空間を制限する。それにより、提案の BF を用いる CI テストの回数の削減を行う。

ベイジアンネットワーク IRT の性質として、真のベイジアンネットワーク IRT はベイジアンネットワークの厳密学習で得られるグラフの部分グラフになることが証明される。この性質を利用し、提案手法では最初にベイジアンネットワークの厳密学習を行う。そして、その結果削除されなかったエッジに対してのみ、提案の BF を用いる CI テストを行う。この方法により、完全グラフから制約ベースアルゴリズムを開始する場合に比べ、提案の BF を用いる CI テストの回数を削減できる。

BDeu は IRT パラメータの推定を必要とせず解析的に求まるため、Natori ら [8] の手法で行う CI テストは提案の BF を用いる CI テストに比べ計算時間が短い。したがって、ベイジアンネットワークの厳密学習に Natori らの手法を用いることで、ベイジアンネットワーク IRT の構造学習全体の時間を短縮できる。実際には、Natori らの手法を高速化した本田ら [13] の手法を用いる。その際、Hashimoto ら [7] と同様に、エッジの向きはテストにおける出題順が前の項目から後の項目であることを仮定する。このベイジアンネットワークの厳密学習を 第 1 段階目の学習と呼ぶ。

提案手法のアルゴリズムを Algorithm3 に示す。

## 5 評価実験

本章では、まず部分グラフ制約により計算コストが削減できることを示すため、提案手法と提案の BF を組み込んだ部分グラフ制約を用いず一般的な制約ベースアルゴリズムを用いて構造学習した場合（以下、制約ベース）の計算時間を比較する。制約ベースは、Algorithm 3 における入力グラフ  $G$  をベイジアンネットワークの厳密学習結果ではなく完全グラフに置き換えたものである。

次に、データ数が十分に大きいときに提案手法が従来手法よりも高精度な学習ができることを示すため、人工的に作成した反応データを用いて提案手法と ILS 分析の学習の精度を比較する。

### 5.1 実験手順

5 項目、10 項目、50 項目からなる 3 種類のテストを 1 万人の被験者に実施した場合を想定した実験を行なった。具体的には、項目数  $J = 5, 10, 50$  の各々について、以下の手順で実験を行なった。ここで、項目間の依存性の強さを表すハイパーパラメータを  $\mu_d$  とする。

---

**Algorithm 3** 提案アルゴリズム

---

**Require:**  $N$  ノード変数, データ  $U$

**Ensure:** ネットワーク構造  $G$

```
1:  $G \leftarrow$  ベイジアンネットワークの厳密学習結果
2:  $t \leftarrow 0$ 
3: while  $t < N - 2$  do
4:   for 各ノード  $X_i$  do
5:     for  $X_i$  の各親ノード  $X_j$  do
6:        $C \leftarrow X_j$  を除く  $X_i$  の親ノード集合中の  $t$  個のノードの組み合わせ
7:       for  $c \in C$  do
8:         if  $CI(X_i, X_j|c, G, U) = 1$  then
9:            $G \leftarrow G$  からエッジ  $(X_i, X_j)$  を削除したグラフ
10:          break : ただし  $CI(X_i, X_j|c, G, U)$  は提案の BF を用いる条件付き独立性
            検定関数.
11:         end if
12:       end for
13:     end for
14:   end for
15:    $t \leftarrow t + 1$ 
16: end while
17: return  $G$ 
```

---

1. 項目数  $J$  の真のベイジアンネットワーク IRT の構造を発生させる.
2. 項目  $J$  個分の項目パラメータと被験者 1 万人分の能力パラメータをランダムに生成する.
3. 1 および 2 で生成したベイジアンネットワーク IRT を用いて, 被験者 1 万人分の反応データを発生させる.
4. 3 で発生させた反応データを用いて, ベイジアンネットワーク IRT の構造学習を行う.
5.  $\mu_d = 1.95, 1.00$  の各々の場合について, 2~4 を行う.
6. 1~5 を 5 回繰り返す.
7. 得られた 5 回分の構造学習の結果から, 手法の評価指標を求める.

以下, 各手順の詳細を述べる.

### 5.1.1 真のベイジアンネットワーク IRT の構造の生成

先行研究 [6] に従い, 真のベイジアンネットワーク IRT の構造において, 各項目は次の 3 種類のいずれかとして生成した.

1. 親項目を持たない
2. 前に出題された項目の中に親項目が 1 つある

3. 前に出題された項目の中に親項目が2つある

(1) の生成確率は 0.3, (2) の生成確率は 0.5, (3) の生成確率は 0.2 とし, これらの確率にしたがって全項目を生成した. (2) および (3) の親項目は, 当該項目より前に出題された項目の中からランダムに選んだ.

### 5.1.2 項目パラメータの真値の設定

識別力パラメータは, 範囲を  $0.3 < a < 2.5$  として以下の対数正規分布から発生させた.

$$\log a \sim N(0, 1)$$

先行研究 [6] に従い, 次のように難易度パラメータを生成した. 親項目を持たない項目の場合には, 難易度パラメータは以下の正規分布から発生させた.

$$b \sim N(0, 1)$$

親項目をもつ項目の場合には, 以下の正規分布に従う確率変数  $b_{j'c}$ ,  $d_{j'}$  を用いて難易度パラメータを生成した.

$$\begin{aligned} b_{j'c} &\sim N(0, 0.25) \\ d_{j'} &\sim N(\mu_d, 0.01) \end{aligned}$$

ここで  $\mu_d$  は依存性の強さを決めるハイパーパラメータであり, 強い依存性のとき 1.95, 弱い依存性のとき 1.0 とした.

項目  $j$  の親項目すべてに正答した時の難易度パラメータは  $b_{j'c} - d_{j'}$  とし, 少なくとも一つの親項目に誤答したときの難易度パラメータは  $b_{j'c} + d_{j'}$  とした.

被験者の能力パラメータを標準正規分布  $N(0, 1)$  から 1 万個発生させた.

### 5.1.3 反応データの生成

発生させた真のベイジアンネットワーク IRT から被験者 1 万人分の反応データを発生させた.

### 5.1.4 構造学習

提案手法および制約ベースの事前設定として, 項目パラメータの MAP 推定における数値積分の求積点数は 15, 収束基準は 0.0001 とし, パラメータの更新前と更新後の変化量がこの値を下回ったときに収束したと判定する. MAP における項目パラメータの事前分布は, 識別力パラメータが  $N(1.0, 0.25^2)$ , 難易度パラメータが  $N(0, 1)$  とした. 提案手法において BIC の計算で仮定する  $\theta$  の事前分布は標準正規分布  $N(0, 1)$  とした.

ベイジアンネットワークの厳密学習に用いる本田ら [13] の構造学習アルゴリズムにおいて, ESS は 1.0 を用いた.

ILS 分析の LCI 検定の閾値は Hashimoto ら [7] が用いている 0.01, 0.05, 0.10 とした.

構造学習の制限時間として 6 時間を設けた. 提案手法および制約ベースの実行環境は, Ubuntu20.04 を OS とする計算機 (Intel(R) Core(TM) i9-10900K 3.70GHz, RAM 64GB) で

ある。ILS 分析の実行環境は Windows10 を OS とする計算機 (Intel(R) Core(TM) i7-10710U 1.10GHz, RAM 16GB) である。制限時間を設けたのは、提案手法が現実時間で学習できることを示すためである。

また、使用するメモリの上限として 60GB を設けた。

### 5.1.5 構造学習手法の評価

各手法の評価指標として、平均計算時間と Structural Hamming Distance (SHD) [14] を用いた。SHD は、真には存在するが学習において削除されたエッジ数、真には存在しないが学習において残されたエッジ数、エッジの方向付けの誤り数の和によって真の構造と推定された構造の距離を表す。SHD はベイジアンネットワークの構造学習の評価に一般的に用いられる指標である。

なお提案手法、制約ベース、ILS 分析ともに真のエッジの向きは既知であることを仮定するため、本実験ではエッジの方向付けの誤り数は SHD にカウントされない。

## 5.2 実験結果

### 5.2.1 提案手法と制約ベースの比較

提案手法と制約ベースの平均計算時間の結果を表 1 に示す。表中の"N/A"はメモリ不足により学習ができなかったことを示す。提案手法の計算時間における () 内の数字は、第 1 段階目の学習であるベイジアンネットワークの厳密学習に要した計算時間を示す。

また、提案手法中のベイジアンネットワークの厳密学習により得られたグラフと完全グラフにおける難易度パラメータ数の平均を表 3、エッジの本数の平均を表 5 に示す。真の構造に含まれるエッジのうちベイジアンネットワークの厳密学習結果のグラフに含まれないエッジの本数の平均を表 7 に示す。

提案手法と制約ベースの平均 SHD の結果を表 9 に示す。

表 1, 2 より、提案手法の計算時間は制約ベースの計算時間より短いことがわかる。よって、第 1 段階目の学習としてベイジアンネットワークの厳密学習を行うことで、計算コストを削減できたことがわかる。

また、表 1, 2 より、項目数が 50 のとき、制約ベースではメモリ不足が発生したのに対し、提案手法ではメモリ不足が発生しなかったことが示されている。したがって、第 1 段階目の学習としてベイジアンネットワークの厳密学習を行うことで、使用するメモリの容量を削減できたと考えられる。

ベイジアンネットワークの厳密学習により使用するメモリの容量を削減できた理由を考察する。項目数が 50 のとき、完全グラフにおける難易度パラメータの数は  $2^{50} - 1$  ( $= 1125899906842623$ ) 個になる。一方、表 3, 4 より、ベイジアンネットワークの厳密学習結果のグラフにおける難易度パラメータの数の平均はおよそ 400~500 個であり、完全グラフの難易度パラメータ数に比べて非常に少ない。このように難易度パラメータの数に大きな差が生じるのは、ベイジアンネットワーク IRT において、ある項目の難易度パラメータの数がその親項目数に応じて指数的に増加するためである。提案の CI テストを行う際には、比較する 2 つの構造の難易度パラメータを全て推定す

る必要がある。したがって、提案手法において推定するパラメータ数は、制約ベースにおいて推定するパラメータ数より非常に少ないため、使用するメモリの容量を削減できたと考えられる。

続いて、ベイジアンネットワークの厳密学習により計算コストを削減できた理由を考察する。1つ目に、構造学習で使用するメモリ容量が削減できた理由と同様に、提案手法では提案の CI テストを行う際に推定するパラメータ数を減らせることが考えられる。表 3, 4 より、いずれの項目数においても、ベイジアンネットワークの厳密学習結果のグラフにおける難易度パラメータの数は完全グラフの難易度パラメータ数より少ないことが読み取れる。よって、第 1 段階目の学習としてベイジアンネットワークの厳密学習を行うことで、第 2 段階目の学習において推定する難易度パラメータの数が減るため、計算コストを削減できると考えられる。2つ目に、ベイジアンネットワーク厳密学習により、提案の CI テストを行うエッジの本数を削減できたことが考えられる。表 5, 6 より、いずれの項目数においても、提案手法において CI テストを行うエッジの平均本数は、制約ベースにおいて CI テストを行うエッジの平均本数よりも少ないことがわかる。したがって、第 1 段階目の学習としてベイジアンネットワークの厳密学習を行うことで、提案の BF を用いる CI テストを行うエッジの本数を削減し、計算コストを削減できると考えられる。

次に、データ数が十分に大きいとき、ベイジアンネットワークの厳密学習により解の探索空間を制限できることを確認する。表 7, 表 8 より、項目数が 5, 10 のとき、ベイジアンネットワークの厳密学習により得られたグラフが、真の構造に含まれるエッジをすべて含んでいることが示されている。さらに表 5, 6 から、ベイジアンネットワークの厳密学習により得られるグラフは、完全グラフに比べてエッジの本数が少ないことがわかる。したがって、項目数に対してデータ数が大きいとき、ベイジアンネットワークの厳密学習を行うことでベイジアンネットワーク IRT の解の探索空間を制限できると考えられる。

データ数が十分でないときのベイジアンネットワークの厳密学習の有効性について検証する。表 7 より、項目数が 50 のとき、真の構造に含まれるエッジのうちベイジアンネットワークの厳密学習結果のグラフが含まないエッジがあったことがわかる。したがって、項目数に対してデータ数が少ないとき、ベイジアンネットワークの厳密学習により真の構造に含まれるエッジが誤って削除されると考えられる。しかし表 1, 2 より、制約ベースではメモリ不足が発生したことが読み取れる。よって、メモリ容量の観点から、データ数が少ない場合であっても、項目数が多い大規模な構造に対しては第 1 段階目の学習を行う提案手法が推奨される。

最後に、提案手法と制約ベースの SHD を比較する。表 9, 10 より、いずれの項目数および依存性の強さにおいても、提案手法は制約ベースと同等以上の精度を有することがわかる。しかし、依存性が強い場合には、項目数 5 の構造における提案手法と制約ベースの SHD が一致している。また依存性が弱い場合であっても、項目数 5 の構造における提案手法と制約ベースの SHD の差は 0.6 であり、項目数 10 の構造における SHD の差より小さいことがわかる。したがって、項目数に対してデータ数が大きいとき、提案手法と制約ベースの精度は同程度になることが示唆される。また、項目数が 10 の構造において、提案手法の方が制約ベースよりも精度が良い。よって、項目数に対してデータ数が十分でないとき、最初にベイジアンネットワークの厳密学習を行うことで、構造学習の精度が向上する可能性がある。

表 1 平均計算時間 [s] (依存性が強い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>145.18</b> (7.84)	<b>474.41</b> (10.43)	<b>11073.88</b> (21.74)
制約ベース	316.14	4164.23	N/A

表 2 平均計算時間 [s] (依存性が弱い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>230.56</b> (8.30)	<b>465.41</b> (7.14)	<b>13302.53</b> (98.60)
制約ベース	459.69	1466.97	N/A

表 3 各グラフにおける難易度パラメータ数の平均 [個] (依存性が強い場合)

手法	項目数 5	項目数 10	項目数 50
ベイジアンネットワークの 厳密学習結果のグラフ	16.2	56.8	379.8
完全グラフ	31.0	1023.0	1125899906842623.0

表 4 各グラフにおける難易度パラメータ数の平均 [個] (依存性が弱い場合)

手法	項目数 5	項目数 10	項目数 50
ベイジアンネットワークの 厳密学習結果のグラフ	19.0	85.4	493.8
完全グラフ	31.0	1023.0	1125899906842623.0

表 5 各グラフにおけるエッジの本数の平均 [本] (依存性が強い場合)

グラフの種類	項目数 5	項目数 10	項目数 50
ベイジアンネットワークの 厳密学習結果のグラフ	7.0	19.4	120.4
完全グラフ	10.0	45.0	1225.0

表 6 各グラフにおけるエッジの本数の平均 [本] (依存性が弱い場合)

グラフの種類	項目数 5	項目数 10	項目数 50
ベイジアンネットワークの 厳密学習結果のグラフ	7.6	24.8	139.6
完全グラフ	10.0	45.0	1225.0

表 7 真の構造に含まれるエッジのうちベイジアンネットワークの厳密学習結果のグラフに含まれないエッジの本数の平均 [本] (依存性が強い場合)

項目数 5	項目数 10	項目数 50
0.0	0.0	1.6

表 8 真の構造に含まれるエッジのうちベイジアンネットワークの厳密学習結果のグラフに含まれないエッジの本数の平均 [本] (依存性が弱い場合)

項目数 5	項目数 10	項目数 50
0.0	0.0	1.8

表 9 提案手法と制約ベースの平均 SHD (依存性が強い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>0.2</b>	<b>0.4</b>	<b>16.8</b>
制約ベース	<b>0.2</b>	3.0	N/A

表 10 提案手法と制約ベースの平均 SHD (依存性が弱い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>1.2</b>	<b>3.8</b>	<b>9.8</b>
制約ベース	1.8	9.6	N/A

## 5.2.2 提案手法と ILS 分析の比較

提案手法と ILS 分析の平均 SHD の結果を表 11, 12 に示す. また, 提案手法と ILS 分析の平均計算時間の結果を表 13, 14 に示す. 提案手法の計算時間における () 内の数字は, 第 1 段階目の学習であるベイジアンネットワークの厳密学習に要した計算時間を示す.

まず, 項目間の依存性の強さが SHD に与える影響を検証する. 表 11 より, 依存性が強いとき, いずれの項目数においても提案手法が最も高精度であった. 一方, 表 12 より, 依存性が弱いとき, 項目数が 10 の構造では ILS 分析 (閾値:0.05) が最も高精度であった. したがって, 項目間の依存性が強いとき, 提案手法は従来手法より高精度な学習ができることがわかる.

続いて, 項目間の依存性の強さと ILS 分析の閾値について考察する. 表 11 より, 項目間の依存性が強いとき, ILS 分析では閾値 0.01 が最も高精度であった. それに対し, 表 12 より, 依存性が弱いとき, ILS 分析の閾値 0.05 が最も高精度であった. よって, ILS 分析で用いる閾値の最適値は, 項目間の依存性の強さによって変わることが示唆される.

次に, 項目数に対するデータ数が SHD に与える影響を検証する. 表 11, 12 より, 項目数が 5 のとき, 依存性の強さに関係なく提案手法が最も高精度であった. すなわち, 項目数に対してデー



タ数が大きいとき、提案手法が最も高精度であった。これは提案手法が漸近一致性をもつためであると考えられる。

また、表 11, 12 より、ILS 分析 (閾値 0.01, 0.05, 0.10) では項目数が多くなるにつれて SHD が著しく増加した。一方、提案手法は項目数の増加とともに SHD が増加するが、従来手法に比べて増加分は少ないことが示されている。CMI を用いる検定では、うまく閾値を選べば少数データであっても精度が良くなると考えられる。したがって、項目数が多い場合には提案手法よりも ILS 分析の方が高精度であると予想されたが、本実験は異なる結果を示した。これには次の 2 つの原因が考えられる。

1 つ目は、提案手法では低次の CI テストを行うのに対し、ILS 分析では低次の CI テストを行わないためである。Algorithm1 にあるように、ILS 分析で用いる LCI 指標は完全グラフを仮定して計算される。このことは、項目  $j$  と  $j'$  の LCI 検定を行う際には、項目 1 から  $j-1$  をすべてを所与とする CI テストを行うことを意味する。すなわち、LCI 検定は常に高次の CI テストを行う。一方、提案手法は一般的な制約ベースアルゴリズムをベースにしており、まず所与とする項目が少ない低次の CI テストを行い、それで削除されなかった場合に高次の CI テストを行う。ベイジアンネットワークの分野では、高次の CI テストはパラメータの収束が遅く、低次の CI テストに比べ信頼性が低いことが知られている。したがって、ILS 分析は項目数が多くなると CI テストの精度が著しく低下すると考えられる。

2 つ目は、提案手法では、条件付き確率にロジスティック関数を仮定するためである。LCI 検定では、条件付き確率を母比率の最尤推定値として求める。一方、提案手法は、ベイジアンネットワーク IRT において、能力  $\theta$  の被験者が項目に正答する条件付き確率が  $\theta$  のロジスティック関数で表現できることを利用し、項目パラメータ  $a, b$  の推定値を求めることで条件付き確率の推定を行う。よって、少数データの場合、条件付き確率に  $\theta$  の関数の形を仮定する提案手法のほうが、従来手法より条件付き確率の推定精度が高いと考えられる。

次に、計算時間を比較する。いずれの手法も、すべての構造学習において制限時間内に学習が終了した。表 13, 14 より、ILS 分析の計算時間は提案手法に比べ非常に短いことがわかる。その理由として、LCI 検定で用いる LCI 指標が提案の BF より短時間で求められることが挙げられる。式 (2) より、LCI 指標はベイジアンネットワーク IRT の項目パラメータの推定を必要とせず、特定の反応パターンをとった被験者数のカウントのみで計算が可能であることがわかる。したがって、項目パラメータを繰り返し計算によって求める提案の BF に比べ、LCI 指標を求める計算は高速に実行できるため、ILS 分析の方が構造学習に要する時間が少ないと考えられる。

表 11 提案手法と ILS 分析の平均 SHD (依存性が強い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>0.2</b>	<b>0.4</b>	<b>16.8</b>
ILS 分析 (閾値 0.01)	1.2	4.4	63.4
ILS 分析 (閾値 0.05)	2.2	5.0	41.8
ILS 分析 (閾値 0.10)	3.4	5.6	42.0

表 12 提案手法と ILS 分析の平均 SHD (依存性が弱い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	<b>1.2</b>	3.8	<b>9.8</b>
ILS 分析 (閾値 0.01)	2.8	2.8	52.8
ILS 分析 (閾値 0.05)	2.2	<b>2.6</b>	42.4
ILS 分析 (閾値 0.10)	3.0	5.0	43.2

表 13 平均計算時間 [s] (依存性が強い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	145.18 (7.84)	474.41 (10.43)	11073.88 (21.74)
ILS 分析 (閾値 0.01)	<b>0.03</b>	0.03	5.74
ILS 分析 (閾値 0.05)	0.05	<b>0.02</b>	6.06
ILS 分析 (閾値 0.10)	<b>0.03</b>	<b>0.02</b>	<b>5.01</b>

表 14 平均計算時間 [s] (依存性が弱い場合)

手法	項目数 5	項目数 10	項目数 50
提案手法	230.56 (8.30)	465.41 (7.14)	13302.53 (98.60)
ILS 分析 (閾値 0.01)	<b>0.03</b>	0.03	<b>9.56</b>
ILS 分析 (閾値 0.05)	0.05	<b>0.02</b>	9.59
ILS 分析 (閾値 0.10)	0.05	0.03	9.85

以上の結果は次のようにまとめられる.

- 真のベイジアンネットワーク IRT がベイジアンネットワークの厳密学習結果のグラフの部分グラフになる制約を用いることで、構造学習の計算コストを削減できた.
- 提案手法により、(1) 項目数に対してデータ数が大きい場合、(2) 項目間の依存性が強い場合に、ILS 分析よりもベイジアンネットワーク IRT の構造学習の精度を改善することができた. 特に (1) は、提案手法の漸近一致性によるものだと考えられる.
- 項目数に対してデータ数が非常に少ない場合、提案手法の方が ILS 分析より高精度であった. その理由として (1) 提案手法は信頼性の高い低次の CI テストを行ってから高次の CI テストを行うこと、(2) 提案手法は条件付き確率に  $\theta$  の関数を仮定するため条件付き確率の推定精度が高いことが考えられる.

## 6 むすび

本論文では、ILS 分析が漸近一致性をもたない問題を改善するため、BIC による BF を用いる CI テストを組み込んだ制約ベースアルゴリズムを提案した。その際、計算コストを削減するため、真のベイジアンネットワーク IRT がベイジアンネットワークの厳密学習により得られるグラフの部分グラフになる制約を用いた。シミュレーションデータによる実験を行なったところ、部分グラフ制約を用いることで、制約を用いない場合に比べ計算コストを削減できた。また提案手法により、(1) 項目数に対してデータ数が大きい場合、(2) 項目間の依存性が強い場合に、ILS 分析よりベイジアンネットワーク IRT の構造学習の精度を向上できた。(1) の場合における精度向上は、提案手法が漸近一致性をもつためであると考えられる。また、項目数に対してデータ数が非常に少ない場合、ILS 分析の方が高精度であると予想されたが、実際には提案手法の方が高精度であった。これは (1) ILS 分析が高次の CI テストのみ行うのに対し、提案手法は低次の CI テストを行ってから高次の CI テストを行う、(2) ILS 分析が条件付き確率そのものをパラメータとして推定するのに対し、提案手法は条件付き確率が  $\theta$  を変数とするロジスティック関数で表せることを利用して項目パラメータを推定するため、条件付き確率の推定精度が高くなることが要因であると考えられる。

今後の課題として、構造学習に要する計算時間の削減が挙げられる。尤度を  $\theta$  について周辺化する提案の BF を用いる場合、いずれの 2 項目間の CI テストを行う際にもベイジアンネットワーク構造全体の BF を求める必要があるため、計算効率が悪い。提案の BF において尤度を  $\theta$  について周辺化しない場合、BF は分解可能になる。この分解可能な BF を用いる場合、検定対象の 2 項目とその親項目の部分グラフのみの BF を求めるだけで CI テストが可能になる。したがって、 $\theta$  を高速かつ高精度に推定できれば、分解可能な BF を用いることで、構造学習の精度を落とさず、提案手法より計算時間を削減できる可能性がある。

## 参考文献

- [1] F.M. Lord, M.R. Novick, and Allan Birnbaum. *Statistical theories of mental test scores*. Addison-Wesley, 1968.
- [2] Wendy M. Yen. Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, Vol. 30, pp. 187–213, 1993.
- [3] C.T Cheng and W.C. Wang. Effects of ignoring item interaction on item parameter estimation and detection of interacting items. *Applied Psychological Measurement*, Vol. 31, pp. 388–411, 2007.
- [4] 登藤直弥. 局所独立性の仮定が満たされない場合の潜在特性推定への影響. 日本テスト学会誌, Vol. 6, pp. 17–28, 2010.
- [5] Maomi Ueno. An extension of the *IRT* to a network model. *Behaviormetrika*, Vol. 29, No. 1, pp. 59–79, 2002.
- [6] 橋本貴充, 植野真臣. 潜在変数周辺化による項目潜在構造分析. 日本教育工学会論文誌, Vol. 35, No. 3, pp. 205–215, 2011.
- [7] Takamitsu HASHIMOTO and Maomi UENO. Latent conditional independence test using bayesian network item response theory. *IEICE Transactions on Information and Systems*, Vol. E94.D, No. 4, pp. 743–753, 2011.
- [8] Kazuki Natori, Masaki Uto, and Maomi Ueno. Consistent learning bayesian networks with thousands of variables. In *Advanced Methodologies for Bayesian Networks (Proc. Machine Learning Research)*, Vol. 73, pp. 57–68, 2017.
- [9] H. Steck and T.S. Jaakkola. On the dirichlet prior and bayesian regularization. *Neural Information Processing Systems (NIPS 2002)*, pp. 697–704, 2002.
- [10] Raanan Yehezkel and Boaz Lerner. Bayesian network structure learning by recursive autonomy identification. *Journal of Machine Learning Research*, Vol. 10, pp. 1527–1570, 2009.
- [11] Peter Spirtes, Clark Glymour, Richard Scheines, and Robert Tillman. Automated search for causal relations: Theory and practice. In *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pp. 467–506. College Publications, 2010.
- [12] J. Abellán, M. Gómez-Olmedo, S. Moral, et al. Some variations on the PC algorithm. *International Conference on Probabilistic Graphical Models*, pp. 1–8, 2006.
- [13] 本田和雅, 名取和樹, 菅原聖太, 磯崎隆司, 植野真臣. 推移性を利用した大規模ベイジアンネットワーク構造学習. 電子情報通信学会論文誌 D, Vol. J102-DTsamardinos2006, No. 12, pp. 796–811, 2019.
- [14] Ioannis Tsamardinos. The max-min hill-climbing bayesian network structure learning algorithm. *Machine Learning*, Vol. 65, No. 1, pp. 31–78, 2006.
- [15] 植野真臣. ベイジアンネットワーク. コロナ社, 2013.

- [16] 橋本貴充. 潜在変数を所与とした条件付き独立性検定とテスト項目分析への応用. 電気通信大学 博士論文, 2011.
- [17] 名取和樹, 宇都雅輝, 植野真臣. Bayes factor を用いた RAI アルゴリズムによる大規模ベイジアンネットワーク学習. 電子情報通信学会論文誌 D, Vol. J101-D, No. 5, pp. 754–768, 2018.
- [18] R. Darrell Bock and Murray Aitkin. Marginal maximum likelihood estimation of item parameters: Application of an em algorithm. *Psychometrika*, Vol. 46, No. 1, pp. 443–459, 1981.
- [19] Luis Enrique Sucar. *Probabilistic Graphical Models: Principles and Applications*. Springer Publishing Company, Incorporated, 2015.
- [20] 杉山将. 統計的機械学習: 生成モデルに基づくパターン認識. Tokyo tech be-text. オーム社, 2009.
- [21] Gideon Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, Vol. 6, No. 2, pp. 461 – 464, 1978.
- [22] 須永潤一. ベイジアンネットワーク IRT に対するベイズ推定法の提案. 電気通信大学 修士論文, 2016.