THE IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS (JAPANESE EDITION)

CIC 電子情報通信学会 Di扁文誌 懒 システム

VOL. J106-D NO. 2 FEBRUARY 2023

本PDFの扱いは、電子情報通信学会著作権規定に従うこと。 なお、本PDFは研究教育目的(非営利)に限り、著者が第三者に直接配布すること ができる。著者以外からの配布は禁じられている。

情報・システムソサイエティ

一般社団法人 電子情報通信学会

THE INFORMATION AND SYSTEMS SOCIETY
THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS

Baves Factor を用いたベイジアンネットワーク IRT の制約ベース学習

青木 健登†a) 菅原 聖太†b) 植野 真臣†㎝

Constraint-Based Learning Bayesian Network IRT Using Bayes Factor Kento AOKI^{†a)}, Shouta SUGAHARAI^{†b)}, and Maomi UENO^{†c)}

あらまし e テスティングは、項目反応理論 (Item Response Theory; IRT) と呼ばれるテスト理論に基づき、テス ト構成、スコア採点が行われる、一般に IRT は、被験者の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、という局所独立性を仮定する、しかし、実際のテストではこの仮定が成り立たないことが多い、 この問題を解決するために、Ueno (2002) は、被験者の能力潜在変数をもつ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 を提案し、その有効性を報告している. しかし、ベイジ アンネットワーク IRT の構造学習については提案されていなかった. そこで Hahimoto ら (2011) は、ベイジアン ネットワーク IRT の構造学習法を提案しているが、漸近一致性がないという問題がある。本論文では、ベイジア ンネットワーク IRT 学習のための BIC (Bayesian Information Criterion) による BF (Bayes Factor) を提案し、それを 用いた条件付き独立性検定(CIテスト)を組み込んだ漸近一致性をもつ構造学習手法を提案する. 提案の BF は 計算コストが大きいため、ベイジアンネットワークの厳密学習で得られるグラフを制約に用いることで、構造学 習の計算コストの削減を行う、シミュレーション実験により、提案手法の有効性を示す。

キーワード e テスティング、項目反応理論、ベイジアンネットワーク、ビッグデータ、機械学習

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[1],[2]. e テスティングは, 項目反応理論 (Item Response Theory: IRT) [3] と呼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) 異なる 項目により構成されるテストを実施しても被験者を同 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

†電気通信大学大学院情報理工学研究科,調布市 Graduate School of Informatics and Engineering, The University of Electro-

Communications, 1-5-1 Chofugaoka, Chofu-shi, 182-8585 Japan

DOI:10.14923/transinfj.2022LEP0007

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[4]. 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[5]や、能力推定値の 標準誤差の増加[6] 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 [7] 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組 み入れた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-parameter logistic model (2PL モデル) に比 べて、未知の項目に対する被験者の反応の予測性能が 高いことが報告されている[7]. しかし、Ueno [7] で はベイジアンネットワーク IRT の構造学習法を提案し ていない。

その後、ベイジアンネットワーク IRT の構造学習法 として、統計的検定を用いて、Hashimoto ら[8] 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:

a) E-mail: aoki@ai.lab.uec.ac.jp

b) E-mail: sugahara@ai.lab.uec.ac.jp

c) E-mail: ueno@ai.lab.uec.ac.jp

LCI 検定)と呼ばれる条件付き独立性検定(Conditional Independence Test: CI テスト)[8] を用いて学習を行う. LCI 検定の精度は設定するしきい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[9] 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[9] 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. Ueno [10] は、ベイジアンネッ トワークにおいて、周辺ゆう度である BDeu (Bayesian Dirichlet equivalence uniform) よりも BIC の方が事前 分布のハイパーパラメータに頑健であることを示して いる. また、ベイジアンネットワーク IRT における 周辺ゆう度の積分は解析的に解けず、パラメータ数の 多さから数値積分は困難である. よって、提案手法で は BIC による BF を用いる. 更に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回当りの 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.

漸近一致性をもつベイジアンネットワークの構造学習を、ベイジアンネットワークの厳密学習という。本論文では、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で得られるグラフの部分グラフになることを証明する。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テストを行うことを提案する。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と期待できる。

以上の文脈より、本論文は以下のベイジアンネットワーク IRT 学習法を提案する.

- (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、以下の二つが挙げられる.

- (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) [3] 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 θ 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-parameter logistic model(2PL モデル)では,能力潜在変数が θ_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.

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

- *a_i*: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の程度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) 異なる項目により構成されるテストを実施しても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 モデルのゆう度関数は次のように定義される.

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

- $U = \{u_{ij}\}\ (i = 1, \dots, I; \ j = 1, \dots, J)$
- $\theta = \{\theta_i\} (i = 1, \dots, I)$
- $a = \{a_j\} (j = 1, \dots, J)$
- $b = \{b_i\} (j = 1, \dots, J)$
- 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、全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 θ 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 [4]. 例えば、次のような場合が挙げられる [6].

• 2項目間に直接的なつながりが存在する

(例 前問に対する回答を使って解くような問題)

• 2項目間に間接的なつながりが存在する

(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 θ 以外の特性が反映される

(例 テスト分析者が仮定した θ 以外の特性 (性別, 国籍) によって正答確率が変化する問題)

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[5] や能力パラメータの推定値の標準誤差[6] 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 [7] 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

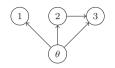


図1 ベイジアンネットワーク IRT の例 Fig. 1 Example of Bayesian network IRT.

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、難易度が変化す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化する。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。

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ゆう度関数は以下の式で表される.

$$L(\boldsymbol{u}_i|\theta_i,\boldsymbol{a},\boldsymbol{b},G)$$

$$= \prod_{j=1}^{J} \prod_{k=1}^{2^{m_j}} \left(\frac{1}{1 + \exp(-1.7a_j(\theta_i - b_{jk}))} \right)^{u_{ij} \cdot u_{ijk}} \cdot \left(1 - \frac{1}{1 + \exp(-1.7a_j(\theta_i - b_{jk}))} \right)^{1 - u_{ij} \cdot u_{ijk}}$$

- $u_i = \{u_{ij}\} (j = 1, \dots, J)$
- $a = \{a_i\} (j = 1, \dots, J)$
- $\boldsymbol{b} = \{b_{ik}\} (j = 1, \dots, J; k = 1, \dots, 2^{m_j})$
- *m_i*: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従って項目 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[7].

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

Hashimoto ら [8] 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)[8] 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を仮定する。 G_c における X_j と $X_{j'}$ の親変数集合を $X^{\gamma j'}$ とする。次に, $X^{\gamma j'}$ 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 $I(X_j, X_{j'}|X^{\gamma j'}, G_c)$ 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。

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

ここで、I は被験者の総数、 N_i は $X^{\neg jj'} = x_i^{\neg jj'}$ という反応パターンを示した被験者の人数、 $N_{x,i}$ は $X_j = x_j$ かつ $X^{\neg jj'} = x_i^{\neg jj'}$ という反応パターンをした被験者の人数、 $N_{x_jx_{j'}i}$ は $X_j = x_j$ 、 $X_{j'} = x_{j'}$ かつ $X^{\neg jj'} = x_i^{\neg jj'}$ 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 指標がしきい値 ε 以上のとき X_j と $X_{j'}$ は能力潜在変数 θ 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 [8] は、しきい値として 0.01, 0.05, 0.1 を用いている。

ILS 分析は、LCI 検定を完全グラフ G_c の全てのエッジ $(X_j, X_{j'})$ (j=1,...,j'-1,j'=2,...,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 Algorithm 1 に示す

しかし、LCI 検定の精度は設定するしきい値に強く 影響を受け、漸近一致性をもたないという問題がある.

Algorithm 1 ILS Analysis

```
Input: N ノード変数. データ U. しきい値 \varepsilon
Output: ベイジアンネットワーク構造 G
 1: Gc ← 完全有向グラフ
 2: for 各ノード X<sub>i'</sub> do
       for G_c における X_{i'} の各親ノード X_j do
          I_{ii'} \leftarrow 項目 j \, \stackrel{.}{\triangleright} \, j' の LCI 指標
 6: end for
 7: G \leftarrow G_c
 8: for 各ノード X<sub>i'</sub> do
       for X_{i'} の各親ノード X_i do
           if I_{ii'} < \varepsilon then
10:
11:
              G \leftarrow G からエッジ (X_i, X_{i'}) を削除したグラフ
13.
       end for
14. end for
15: return G
```

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 (Bayes Factor) を用いた学習、の二つの学習から構成される。

4.1 BIC によるベイジアンネットワーク **IRT** の **BF**

ベイジアンネットワークの分野では、Natori ら [9] は 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.

BF は、二つのモデルの周辺ゆう度の比で表され、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 [11] は、2 ノード間が独立・従属なモデルの周辺ゆう度の比による 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 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 BF(X,Y|Z) とし、その対数である対数 BF を $\log BF(X,Y|Z)$ とすると

$$\log \mathrm{BF}(X, Y | \mathbf{Z}) = \log P(\mathbf{D} | G_1, \alpha) - \log P(\mathbf{D} | G_2, \alpha)$$

(3)

と表される。ここで、 $P(D|G,\alpha)$ はデータを D、ハイパーパラメータを α としたときのベイジアンネットワーク構造 G の周辺ゆう度である。Steck ら [11] では、 $P(D|G,\alpha)$ 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 α を Equivalent Sample Size (ESS)という。Steck らは、式 (3) が 0 以上ならば G_1 、それ以外の場合 G_2 の構造を選択する CI テストを提案した。

Natori ら [9] 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 [12] 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 θ が含まれる。そのため、Natori ら [9] 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。

以下、新たに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 a、難易度パラメータ b をまとめて c 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 c の周辺ゆう度 c d d は次式で表される。

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

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

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

 $f(\xi)$ 、 $f(\theta)$ はそれぞれ項目パラメータ ξ 、能力パラメータ θ の事前分布である。周辺ゆう度 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。

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

Ueno [10] は、ベイジアンネットワークにおいて、周 辺ゆう度である BDeu よりも BIC の方が事前分布のハ イパーパラメータに頑健であることを示している.

また、式 (4) の周辺ゆう度は解析的に解くことができない。 モンテカルロ法などの数値積分を適用する場合でも、周辺ゆう度はパラメータ a,b,θ 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) の周辺ゆう度の対数を 近似する、まず

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

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.

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

ここで、 $H_{\pmb{\xi}_{\mathrm{MAP}}}$ は $\log g(U, \pmb{\xi}, \pmb{\theta}|G)$ の $\pmb{\xi}_{\mathrm{MAP}}$ 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 $\pmb{\theta}$ で積分し、対数をとると

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

を得る. 式 (5) は. サンプル数 N が十分大きいとき

$$\begin{split} & \log \iint g(U, \xi, \theta | G) d\xi d\theta \\ & \simeq \log \int L(U|G, \xi_{\text{MAP}}, \theta) f(\theta) d\theta - \frac{k}{2} \log N \end{split}$$

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

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

と表す。 θ 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。 $\mathbf{\xi}_{\text{MAP}}$ は、 θ 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.

$$\log \operatorname{BF}(G_1, G_2) \simeq \operatorname{BIC}(G_1) - \operatorname{BIC}(G_2) \tag{7}$$

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 $Z \ge \theta$ を所与とした 2 項目 X_i と X_i の条件付き独立性を検定する方法を

Algorithm 2 proposed CI test function

Input: 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

Output: $\theta \ \ \, c \ \,$ 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

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[8]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) $\log BF(G_1, G_2) > 0$ ならば θ 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 θ 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 θ について周辺化したゆう度の対数 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 Algorithm 2 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 [13] に組み込む. これを第2 段階目の学習と呼ぶ.

4.2 ベイジアンネットワーク学習による枝刈り

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回当りの 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 において次の定理が成り立つ.

[定理 4.1] 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.

[証明] 背理法を用いて証明する. 変数集合 Z を所与として Z 変数 Z と Z が独立であることを Z が独立であることを Z と表す. また, 変数集合を Z で表し, 変数 Z と Z の ノード間のエッジを Z と表す.

真のベイジアンネットワーク IRT のグラフ G_{BNIRT} が漸近的にベイジアンネットワークの厳密学習で得られるグラフ G_{BN} の部分グラフにならないことを仮定する. このとき、 G_{BNIRT} 、 G_{BN} のエッジの集合をそれぞれ E_{BNIRT} 、 E_{BN} とすると、ある $X,Y \in X$ について、以下が成り立つ.

$(X,Y) \notin E_{BN} \wedge (X,Y) \in E_{BNIRT}$

 $(X,Y) \notin E_{\text{BN}}$ であるから、ある変数集合 $Z \subseteq X/\{X,Y\}$ について以下が成り立つ [14].

$$I(X,Y|\mathbf{Z}) \tag{8}$$

また、ベイジアンネットワーク IRT の仮定から θ が親をもたないため、以下が成り立つ[15].

$$\neg I(X, Y|Z, \theta) \Rightarrow \neg I(X, Y|Z)$$

対偶をとると次式を得る.

$$I(X, Y|\mathbf{Z}) \Rightarrow I(X, Y|\mathbf{Z}, \theta)$$
 (9)

式 (8) と式 (9) から以下が成り立つ.

$$I(X,Y|\mathbf{Z},\theta) \tag{10}$$

一方、 $(X,Y) \in E_{\text{BNIRT}}$ より、任意の変数集合 $\mathbf{Z}' \subseteq X/\{X,Y\}$ について以下が成り立つ.

 $\neg I(X,Y|\mathbf{Z'},\theta)$

これは式 (10) と矛盾する.

定理 4.1. 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い、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

П

Algorithm 3 proposed Algorithm

```
Input: N ノード変数. データ U
Output: ネットワーク構造 G
 1: G ← 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
      for 各ノード X<sub>i</sub> do
4.
         for X_i の各親ノード X_i do
5:
            C \leftarrow X_i 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
               if CI(X_i, X_j | c, G, U) = 1 then
 8:
                   G \leftarrow G からエッジ (X_i, X_i) を削除したグ
 9:
   ラフ
10:
                   break: tricklim CI(X_i, X_i | c, G, U) it proposed
   CI test function
                end if
11.
12:
             end for
         end for
13:
14:
      end for
15:
      t \leftarrow t + 1
16: end while
17: return G
```

る CI テストの回数を削減できると期待できる.

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 [9] 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 の構造学習全体の計算時間を短縮できる。実際には Natori らの手法を高速化した本田ら [16]の手法を用いる。その際、Hashimoto ら [8] 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。

第1段階目の学習と第2段階目の学習から構成される。提案手法のアルゴリズムを Algorithm 3 に示す.

提案手法と ILS 分析のアルゴリズムの主な相違点は 次の 2 点である.

- 提案手法は CI テストに式 (7) で提案した BF を 用いる. ILS 分析は CI テストに LCI 指標を用いる.
- 提案手法では提案の枝刈りを行う. 従来手法では枝刈りを行わない.

結果として, 従来手法は漸近一致性をもたないのに対して. 提案手法はそれをもつという利点がある.

5. 評価実験

本章では、まず 4.2 で提案された枝刈りにより計算

コストが削減できることを示すため、提案手法と枝刈りを用いず 4.1 の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」と呼ぶ)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

TOEIC や情報処理技術者試験などの e テスティングで実施される試験では被験者が数万人規模になり得る. 一方で,大学の講義の試験では被験者が 100 人程度の場合が考えられる.

また, 先行研究 [8], [17] ではデータ数 10,000 のときの LCI 検定, ILS 分析の有効性が報告されている. ベイジアンネットワーク分野では, 名取ら [18] や Natoriら [9] により, CI テストに Bayes Factor を用いる手法と条件付き相互情報量を用いる手法の構造学習精度の比較実験が行われている. 実験において, 名取ら [18] は5 変数の小規模ネットワークでデータ数 100~10,000, Natoriら [9] は10 変数以上のネットワークでデータ数10,000 以上を設定している.

そこで本論文では 5 項目,10 項目,50 項目からなる 3 種類のテストを 100,500,1,000,5,000,10,000,20,000 人の被験者に実施した場合を想定した実験を行った. 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 μ_d とする.

- (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 100~20,000 人分の能力パラメータをランダムに生成する.
- (3) 1及び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 $100\sim20,000$ 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, $2\sim4$ を行う.
 - (6) 1~5を10回繰り返す.
 - (7) 得られた 10 回分の構造学習の結果から、手

表1 平均 SHD Table 1 Average of SHD.

Sample			$u_d = 1.95$		$\mu_d = 1.00$							
size	担安		ILS	ILS	ILS	提案	制約ベース	ILS	ILS	ILS		
	1处米	叫小小一人	(0.01)	(0.05)	(0.10)	1)比米	叫称,	(0.01)	(0.05)	(0.10)		
					J	= 5						
100	1.6	1.9	2.8	2.7	2.5	1.6	1.9	4.2	2.0	2.0		
100	(1.50)	(1.81)	(1.25)	(1.19)	(1.28)	(1.11)	(1.04)	(1.89)	(1.67)	(1.48)		
500	1.0	1.4	2.7	2.3	2.7	0.6	1.5	2.8	1.6	1.7		
300	(1.10)	(1.36)	(1.19)	(1.42)	(1.19)	(0.66)	(1.20)	(0.98)	(1.02)	(1.27)		
1.000	0.8	0.9	2.3	2.4	2.5	0.5	0.6	2.5	1.3	1.9		
1,000	(0.87)	(0.94)	(1.27)	(1.43)	(1.28)	(0.81)	(0.80)	(1.12)	(1.27)	(1.37)		
5,000	0.5	0.6	1.8	2.2	2.5	0.5	0.2	1.8	1.4	1.9		
3,000	(0.67)	(0.80)	(0.87)	(1.33)	(1.28)	(1.02)	(0.60)	(1.40)	(1.28)	(1.30)		
10.000	0.6	0.6	2.0	2.2	2.5	0.3	0.2	1.9	1.3	2.0		
Size 提:	(0.92)	(0.92)	(1.00)	(1.33)	(1.28)	(0.64)	(0.60)	(1.45)	(1.19)	(1.26)		
20,000 (0. 100 3 (1. 500 1,000 1. 1,000 1. 5,000 1	0.3	0.4	1.9	2.2	2.5	0.2	0.3	2.0	1.4	1.9		
20,000	(0.46)	(0.49)	(1.04)	(2.50)	(1.28)	(0.60)	(0.64)	(1.41)	(1.28)	(1.37)		
	J = 10											
100	3.0	4.6	7.8	6.8	7.2	4.2	4.5	10.5	6.1	6.2		
	(1.73)	(2.62)	(2.32)	(1.94)	(1.94)	(1.78)	(2.77)	(5.18)	(1.97)	(1.99)		
500	1.8	3.8	6.1	6.7	6.9	1.9	3.6	9.1	4.9	5.9		
500	(2.23)	(2.68)	(2.51)	(1.62)	(1.58)	(1.76)	(2.87)	(6.01)	(1.87)	(2.02)		
1,000	1.7	5.3	5.3	6.4	6.8	0.8	3.4	4.6	6.0	6.1		
	(1.68)	(2.33)	(2.233)	(1.43)	(1.54)	(0.63)	(2.76)	(6.07)	(2.33)	(1.73)		
5.000	1.8	5.3	4.3	6.3	6.8	2.0	3.8	5.1	4.9	6.1		
3,000	(1.78)	(3.58)	(1.95)	(2.05)	(2.27)	(1.79)	(1.89)	(3.33)	(1.76)	(2.02)		
5,000	1.1	5.9	3.5	4.2	6.7	0.8	6.6	3.7	4.4	6.1		
	(1.14)	(3.11)	(2.01)	(1.83)	(1.90)	(1.17)	(3.50)	(2.00)	(1.80)	(1.81)		
20,000	0.5	4.5	4.9	6.3	7.0	0.8	8.3	3.8	4.4	6.0		
20,000	(0.67)	(2.42)	(2.34)	(1.80)	(1.61)	(1.08)	(3.87)	(1.60)	(2.11)	(1.90)		
					J =	= 50						
100	29.7	N/A	65.7	44.6	42.2	32.7	N/A	58.7	44.7	44.5		
100	(7.28)		(9.73)	(1.62)	(2.19)	(8.17)		(10.42)	(4.62)	(5.31)		
500	16.3	N/A	70.4	42.5	41.4	19.1	N/A	61.1	43.0	43.1		
300	(8.25)	IN/A	(8.59)	(1.47)	(1.72)	(6.20)	N/A	(9.54)	(6.77)	(5.37)		
1.000	11.7	N/A	70.3	41.5	41.0	17.8	N/A	64.2	41.5	42.8		
1,000	(5.62)	IV/A	(10.48)	(1.83)	(1.62)	(8.70)	IVA	(10.07)	(5.75)	(5.57)		
5,000	7.5	N/A	70.5	38.3	39.7	9.3	N/A	68.1	40.6	42.1		
3,000	(3.56)	11/71	(8.52)	(1.90)	(1.47)	(4.24)	IN/A	(12.64)	(5.49)	(6.18)		
10.000	6.5	N/A	70.9	38.4	39.5	9.6	N/A	65.0	40.3	41.7		
100 (0.00	(3.14)	11/71	(8.55)	(1.90)	(1.62)	(5.33)	IN/A	(12.86)	(5.46)	(5.82)		
20,000	7.0	N/A	67.9	38.3	39.4	6.3	N/A	65.7	38.3	41.6		
20,000	(4.31)	11/71	(10.63)	(1.60)	(3.29)	(3.61)	11/14	(13.68)	(5.76)	(6.11)		

法の評価指標を求める.

(1) では、先行研究[17]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. 親項目をもたない
- 2. 前に出題された項目の中に親項目が一つある
- 3. 前に出題された項目の中に親項目が二つある 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, これらの確率に従って全項目を生成した. (2) 及び(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.

(2) において、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.

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

更に, 先行研究[17] に従い, 次のように難易度パラメータを生成した. 親項目をも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

	Sample	J :	= 5	J =	= 10	J = 50			
	size	$\mu_d = 1.95$	$\mu_d = 1.00$	$\mu_d = 1.95$	$\mu_d = 1.00$	$\mu_d = 1.95$	$\mu_d = 1.00$		
	100	14.2 (6.01)	11.1 (2.07)	26.4 (2.37)	28.4 (4.43)	200.6 (68.82)	216.7 (62.57)		
	500	11.9 (2.21)	13.2 (2.32)	29.7 (4.92)	31.3 (6.00)	198.1 (18.99)	195.3 (11.17)		
難易度	1,000	12.4 (1.85)	14.5 (2.80)	30.7 (5.57)	33.6 (6.70)	217.3 (10.07)	226.1 (32.36)		
無勿及 パラメータ数	5,000	15.8 (3.34)	19.6 (4.98)	42.6 (7.28)	49.6 (13.68)	322.9 (36.06)	327.5 (46.09)		
ハラメータ奴	10,000	17.5 (4.03)	20.4 (4.48)	48.0 (16.69)	55.6 (7.28)	403.0 (44.98)	411.2 (61.23)		
	20,000	17.9 (4.83)	20.0 (4.84)	51.1 (13.42)	76.9 (26.35)	513.5 (57.64)	531.2 (87.45)		
	完全グラフ	$\frac{1}{2^5} - \frac{1}{1} - \frac{1}{1} - \frac{1}{1} - \frac{1}{1} - \frac{1}{1} = \frac{1}{1} - $		2 ^{T0}		$\frac{1}{2} - \frac{1}{2} - \frac{1}$			
	100	5.9 (1.58)	4.8 (0.87)	12.3 (1.10)	12.9 (2.02)	81.6 (7.43)	82.9 (4.21)		
	500	5.3 (0.90)	5.7 (0.90)	13.3 (1.49)	13.9 (2.55)	88.7 (6.40)	87.5 (3.44)		
グラフに存在	1,000	5.4 (0.66)	6.6 (1.10)	13.8 (1.72)	14.9 (2.26)	94.9 (3.33)	93.3 (3.95)		
するエッジ数	5,000	6.6 (1.11)	7.8 (1.33)	17.1 (1.92)	18.9 (2.84)	115.6 (8.92)	114.9 (8.93)		
りるエック奴	10,000	7.2 (1.33)	8.0 (1.18)	17.5 (3.96)	20.3 (1.19)	125.1 (8.09)	125.0 (9.37)		
	20,000	7.3 (1.35)	7.9 (1.22)	18.7 (2.49)	20.3 (3.61)	137.4 (8.97)	138.9 (11.12)		
	完全グラフ	i	0	4	5	12	25		
	100	0.4 (0.66)	0.1 (0.30)	0.8 (0.75)	1.1 (1.22)	11.6 (4.22)	13.0 (4.80)		
	500	0.2 (0.40)	0.0 (0.00)	0.7 (0.90)	0.7 (1.00)	7.8 (3.31)	9.7 (3.69)		
missina	1,000	0.3 (0.46)	0.0 (0.00)	0.8 (0.87)	0.4 (0.49)	5.1 (2.26)	8.3 (2.24)		
missing	5,000	0.0 (0.00)	0.0 (0.00)	0.6 (0.60)	0.5 (0.50)	2.7 (1.68)	3.2 (1.47)		
edge	10,000	0.0 (0.00)	0.0 (0.00)	0.2 (0.40)	0.3 (0.46)	2.4 (1.43)	3.3 (2.00)		
	20,000	0.0 (0.00)	0.0 (0.00)	0.2 (0.40)	0.2 (0.40)	2.6 (1.96)	2.2 (1.25)		
	「完全グラフ」	(5	(5	0			

表 2 難易度パラメータ数及びエッジ数の平均 Table 2 Average of the number of difficulty parameters and edges.

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.

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

ここで μ_d は依存性の強さを決めるハイパーパラメータである. Hashimoto ら [8] に従い, μ_d は強い依存性のとき 1.95、弱い依存性のとき 1.0 とした.

項目jの親項目全てに正答したとき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.

(4) では、提案手法及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、識別力が $\log a \sim N(0,1)$ 、難易度が $b \sim N(0,1)$ とした。提案手法において BIC の計算で仮定する θ 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 [16] 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 検定のしきい値は Hashimotoら [8] が用いている 0.01、0.05、0.10 とした。提案手法及び制約ベースの実行環境は、Ubuntu 20.04 を OSとする計算機 (Intel(R) Core(TM) i9-10900 K 3.70 GHz。

RAM 64 GB) である. ILS 分析の実行環境は Windows 10 を OS とする計算機 (Intel(R) Core(TM) i7-10710 U 1.10 GHz, RAM 16 GB) である. また, 使用するメモリの上限として 60 GB を設けた.

(7)では、各手法の評価指標として平均計算時間と Structural Hamming Distance (SHD)[19]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。

5.2 実験結果

平均 SHD の結果を表 1 に示す. ただし、表中の「制約ベース」は本手法で枝刈りを用いない手法を示し、「ILS(0.01)」、「ILS(0.05)」、「ILS(0.10)」はそれぞれ ILS分析における LCI 検定のしきい値を 0.01, 0.05, 0.10と設定したものを示す. 表中の「N/A」はメモリ不足で学習できなかったことを示す. いずれの従属性の強さ、項目数、データ数においても提案手法は ILS分析よりも高精度であることが分かる.

項目数によらず、提案手法の SHD はデータ数が増えると単調減少している. これは提案の BF を用いる検定が漸近一致性をもつためである. 制約ベースの

表 3 平均計算時間 [s]
Table 3 Average of running time[s].

Sample			$\mu_d = 1.9$			$\mu_d = 1.00$						
size		是案	制約	ILS	ILS	ILS		是案	制約	ILS	ILS	ILS
	1 段階	全体	ベース	(0.01)	(0.05)	(0.10)	1 段階	全体	ベース	(0.01)	(0.05)	(0.10)
						J :						
100	6.41	10.92	4.83	0.01	0.01	0.01	6.41	10.78	4.77	0.01	0.01	0.01
	(0.02)	(0.21)	(0.11)	(0.00)	(0.00)	(0.00)	(0.03)	(0.08)	(0.06)	(0.00)	(0.00)	(0.00)
500	6.42	12.95	9.41	0.01	0.01	0.01	6.41	12.74	8.57	0.01	0.01	0.01
	(0.03)	(1.12)	(1.19)	(0.01)	(0.01)	(0.01)	(0.02)	(0.51)	(0.66)	(0.01)	(0.01)	(0.01)
1,000	6.40	16.02	16.02	0.01	0.01	0.01	6.42	16.93	15.59	0.01	0.01	0.01
	(0.02)	(3.16)	(3.05)	(0.00)	(0.00)	(0.00)	(0.04)	(1.44)	(1.93)	(0.00)	(0.00)	(0.00)
5,000	6.46	65.16	124.09	0.01	0.01	0.01	6.45	85.14	133.17	0.01	0.01	0.01
	(0.02)	(20.47)	(38.57)	(0.00)	(0.00)	(0.00)	(0.04)	(25.70)	(17.59)	(0.00)	(0.00)	(0.00)
10,000	6.49	173.72	282.67	0.02	0.02	0.02	6.47	199.86	320.32	0.01	0.01	0.01
	(0.03)	(77.27)	(78.35)	(0.01)	(0.01)	(0.01)	(0.03)	(38.80)	(32.99)	(0.00)	(0.00)	(0.00)
20,000	6.49	385.45	639.35	0.02	0.02	0.02	6.50	437.25	752.62	0.01	0.01	0.01
	(0.04)	(117.37)	(116.76)	(0.00)	(0.00)	(0.00)	(0.03)	(131.32)	(92.28)	(0.00)	(0.00)	(0.00)
	J = 10											
100	6.65	12.56	27.99	0.01	0.01	0.01	6.65	14.15	35.22	0.01	0.01	0.01
	(0.05)	(0.36)	(40.40)	(0.00)	(0.00)	(0.00)	(0.09)	(4.60)	(38.96)	(0.00)	(0.00)	(0.00)
500	6.64	20.19	420.40	0.02	0.02	0.02	6.55	21.85	556.02	0.01	0.01	0.01
	(0.02)	(2.96)	(980.95)	(0.00)	(0.00)	(0.00)	(0.05)	(5.11)	(481.29)	(0.00)	(0.00)	(0.00)
1,000	6.56	32.91	2691.04	0.01	0.01	0.01	6.57	35.73	1557.49	0.01	0.02	0.02
	(0.07)	(6.02)	(4823.63)	(0.00)	(0.00)	(0.00)	(0.09)	(10.90)	(1598.82)	(0.00)	(0.01)	(0.00)
5,000	7.23	240.64	2849.21	0.01	0.02	0.02	6.96	247.63	4916.36	0.02	0.02	0.02
	(0.58)	(108.25)	(1565.24)	(0.00)	(0.00)	(0.00)	(0.43)	(71.63)	(4986.06)	(0.00)	(0.00)	(0.00)
10,000	7.39	423.33	8415.39	0.03	0.02	0.02	7.32	623.54	5619.80	0.02	0.02	0.02
	(0.93)	(137.80)	(7625.80)	(0.01)	(0.01)	(0.01)	(0.59)	(222.22)	(6378.11)	(0.00)	(0.00)	(0.00)
20,000	7.65	1553.18	10840.28	0.03	0.03	0.03	7.78	2025.46	8798.80	0.03	0.03	0.03
	(0.90)	(742.11)	(9292.09)	(0.01)	(0.01)	(0.01)	(0.96)	(695.69)	(6534.98)	(0.01)	(0.01)	(0.01)
	(0.50)	(, ,=,,,)	(>=>=)	(0.01)	(0.01)	` /	: 50	(0,010,)	(******)	(0.01)	(0.01)	(0.00-)
100	7.31	80.45	N/A	0.14	0.13	0.13	7.10	108.98	N/A	0.15	0.14	0.13
	(0.19)	(16.03)		(0.02)	(0.01)	(0.01)	(0.12)	(51.117)		(0.05)	(0.03)	(0.02)
	7.52	486.95	N/A	0.51	0.51	0.49	7.68	403.12	N/A	0.49	0.50	0.46
500	(0.14)	(278.99)		(0.08)	(0.06)	(0.04)	(0.22)	(99.02)		(0.09)	(0.09)	(0.07)
1,000	8.25	906.03	N/A	0.94	0.93	0.96	8.38	977.81	N/A	0.94	0.91	0.93
	(0.29)	(149.20)		(0.16)	(0.12)	(0.15)	(0.52)	(299.68)		(0.31)	(0.20)	(0.22)
5,000	12.44	6380.48	N/A	3.45	3.43	3.44	13.09	5773.27	N/A N/A	3.18	3.26	3.51
	(1.39)	(1352.44)		(0.78)	(0.88)	(0.64)	(2.68)	(2067.65)		(0.90)	(0.85)	(1.23)
	24.54	14899.54		5.93	6.26	5.96	25.86	14477.51		5.77	5.96	5.89
10,000												
	(4.47)	(3174.45)		(1.16)	(1.48)	(1.22)	(9.12)	(2021.13)		(1.97)	(1.88)	(1.93)
20,000	70.13	38911.19	N/A	10.38	10.24	10.62	74.56	37432.71	N/A	11.76	10.43	10.46
	(23.13)	(7510.74)		(2.71)	(2.23)	(2.17)	(39.47)	(7093.78)		(5.36)	(3.68)	(3.63)

SHD は安定しない. ILS(0.01) と ILS(0.05) の SHD は $1.3\sim2.0$ 付近までしか減少せず、ILS(0.10) の SHD は ほぼ一定の高い値を示した. これは ILS 分析が漸近一 致性をもたず、設定するしきい値に強く影響を受ける ためである.

制約ベースは、項目数5のとき提案手法と同等の精度を示したが、項目数が増えると精度が著しく悪化している。第1段階目の学習結果及び完全グラフの難易度パラメータ数の平均とエッジ数の平均を表2に示す。表2より、項目数が増えると、第1段階目の学習結果に比べて完全グラフの難易度パラメータは著しく増加

することが分かる。すなわち制約ベースでは、項目数の増加に伴い推定すべきパラメータ数が非常に大きくなるため、一つのパラメータ当りのデータ数が少なくなってしまい、検定の精度が低くなってしまう。一方、提案手法では枝刈りにより推定すべきパラメータ数が抑えられ、制約ベースほど学習精度が悪化しない。制約ベースが項目数50の場合にメモリ不足で計算できなかったこともパラメータ数の増加が原因である。

次に、ベイジアンネットワークの厳密学習による CI テスト数削減が有効に働いたかについて検証する. 表 2 において、「missing edge」は真の構造に存在するエッ

ジのうちグラフに存在しないエッジ数を示す.表2より,項目数10,50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るグラフは,完全グラフに比べてエッジ数が大きく減少していることが分かる.更に、いずれの項目数においても,データ数が増加するにつれて missing edge が減少することも確認される.これらより,ベイジアンネットワークの厳密学習により枝刈りが行え,提案のCIテスト数を削減し,計算コストを削減できることが分かる.

最後に、各手法の計算時間を表3で比較する、表中 の「1段階」は第1段階目の学習、「全体」は提案手法 全体の計算時間を示す. 表 3 より. 項目数 10.50 の とき、提案手法は制約ベースより計算時間が短い. し たがって、項目数10以上では、第1段階目のベイジ 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示された。また項目数5のとき でも、データ数 5.000 以上の場合は提案手法の方が制 約ベースよりも高速であった. これはデータ数が増え ると提案の BF の計算コストが増加するためである. よって、項目数5の小規模ネットワークであっても、 データ数が増加すると提案の枝刈りにより計算コスト が削減できる.一方、ILS 分析の計算時間は提案手法 に比べ非常に短い. 式(2)より, LCI 指標は特定の反 応パターンをとった被験者の人数のカウントのみで計 算が可能である. 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 の計算は高速に実行できるため、ILS分析の方が計算 時間が短い.

しかし、項目数が増加すると全ての項目反応パターンには欠測データが増加してしまい、ILS 分析の学習精度は大きく劣化してしまう。対照的に提案手法では、枝刈りにより学習精度は大きく劣化しない。

以上より以下の結果を得た.

- シミュレーション実験において、提案手法は ILS 分析よりも高精度であった。
- 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.

6. t t t

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。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て、枝刈りを行った。シミュレーションデータによる実験を行ったところ、枝刈りを行うことで、枝刈りを行わ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、ILS 分析に比べ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。

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。提案のBFは分解可能でないため、CIテストごとに構造全体のBICを求める必要があり、計算効率が悪い。分解可能なスコアを用いる場合、検定対象の2項目とその親項目のみからなる部分グラフのスコアを求めるだけでCIテストが可能になる。したがって、提案手法より計算時間を削減できる可能性がある。また、Tsutsumiら[20],[21]は、Deep-IRTが従来のIRTに比べ、未知の項目に対する被験者の反応の予測精度が高いことを報告している。Bayesian network IRTと Deep-IRT の性質の違いを比較することも今後の課題とする.

対 対

- M. Ueno, K. Fuchimoto, and E. Tsutsumi, "E-testing from artificial intelligence approach," Behaviormetrika, vol.48, no.2, pp.409– 424, 2021.
- [2] M. Ueno, "Ai based e-testing as a common yardstick for measuring human abilities," 2021 18th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE), pp.1–5, 2021.
- [3] F.M. Lord, M.R. Novick, and A. Birnbaum, Statistical theories of mental test scores., Addison-Wesley, 1968.
- [4] W.M. Yen, "Scaling performance assessments: Strategies for managing local item dependence," J. Educational Measurement, vol.30, pp.187–213, 1993.
- [5] 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.
- [6] 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.
- [7] M. Ueno, "An extension of the IRT to a network model," Behaviormetrika, vol.29, no.1, pp.59–79, 2002.
- [8] T. Hashimoto and M. Ueno, "Latent conditional independence test using bayesian network item response theory," IEICE Trans. Inf. & Syst., vol.E94-D, no.4, pp.743-753, April 2011.
- [9] K. Natori, M. Uto, and M. Ueno, "Consistent learning bayesian networks with thousands of variables," Advanced Methodologies for Bayesian Networks (Proc. Mach. Learn. Res.), vol.73, pp.57– 68, 2017
- [10] M. Ueno, "Robust learning bayesian networks for prior belief," UAI 2011, Proc. 27th Conf. Uncertainty Artif. Intell., pp.698–707, AUAI Press, 2011.

- [11] H. Steck and T.S. Jaakkola, "On the dirichlet prior and bayesian regularization," Neural Information Processing Systems (NIPS 2002), pp.697–704, 2002.
- [12] R. Yehezkel and B. Lerner, "Bayesian network structure learning by recursive autonomy identification," J. Mach. Learn. Res., vol.10, pp.1527–1570, 2009.
- [13] P. Spirtes, C. Glymour, R. Scheines, and R. Tillman, "Automated search for causal relations: Theory and practice," Heuristics, Probability and Causality: A Tribute to Judea Pearl, pp.467–506, College Publications, 2010.
- [14] P. Spirtes, C. Glymour, and R. Scheines, Causation, Prediction, and Search, vol.81, Springer, 1993.
- [15] D. Koller and N. Friedman, Probabilistic Graphical Models: Principles and Techniques, Adaptive computation and machine learning, MIT Press, 2009.
- [16] 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ペイジアンネットワーク構造学習," 信学論(D), vol.J102-D, no.12, pp.796-811, Dec. 2019.
- [17] 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,
- [18] 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ペイジアンネットワーク学習," 信学論(D), vol.J101-D, no.5, pp.754-768, May 2018.
- [19] I. Tsamardinos, "The max-min hill-climbing bayesian network structure learning algorithm," Mach. Learn., vol.65, no.1, pp.31– 78, 2006.
- [20] E. Tsutsumi, R. Kinoshita, and M. Ueno, "Deep item response theory as a novel test theory based on deep learning," Electronics, vol.10, issue.9, no.1020, 2021.
- [21] E. Tsutsumi, R. Kinoshita, and M. Ueno, "Deep-IRT with independent student and item networks," EDM, 2021.

(2022 年 3 月 21 日受付, 8 月 4 日再受付, 10 月 25 日早期公開)



青木 健登

2022 電気通信大学情報理工学域卒. 同年, 同大学院情報理工学研究科情報・ネットワーク工学専攻博士前期課程入学, 現在に至る.



菅原 聖太

2020 電気通信大学大学院情報理工学研究 科博士前期課程了.同年,電気通信大学大 学院情報理工学研究科博士後期課程入学, 在学中.



植野 真臣 (正員)

1992 神戸大学大学院教育学研究科了, 1994 東京工業大学大学院総合理工学研究科了, 博士(工学). 東京工業大学, 千葉大学, 長岡技術科学大学を経て 2006 より電気通信大学助教授, 2013 より教授, 現在に至る.