

データマイニングと機械学習による 学力テストの問題間の関係性抽出の試み

Examination on Inter-Item Dynamics in Achievement Tests
through Data-Mining and Machine-Learning Techniques

元木崇¹・佐藤賢二^{1,#}・谷内通¹ (¹金沢大学)

Takashi MOTOKI¹ , Kenji SATO^{1,#}and Tohru TANIUCHI¹ (¹Kanazawa Univ.)



ネットワーク図の
htmlファイル等を
閲覧できます。

本研究の立ち位置

従来

García et al.(2011)は教育現場でも運用できる**Association Rule(AR)**
抽出と共有の枠組みを確立。共起の“見える化”と運用性が強み。
Shimizu et al. (2006) は **LiNGAM** により**非ガウス性**を手掛かりに**因果**
方向を推定する一般的アルゴリズムを提示。

本研究

問題間に焦点化し、**AR(共起)**と**LiNGAM(因果)**を同一データで統合。
能力残差化とブートストラップCIによる**頑健な方向推定**、**群間比較**ま
で一気通貫で行った。

学習行動や
教材利用ログ

設問×設問の微視的
ネットワーク

データセット・方法

データ

- 文科省「全国学力・学習状況調査」
算数(パブリックユースデータ)
- 受検者×設問の反応行列
設問→上位ドメインを対応付け。
- 能力分位で G1 (高) ~G4 (低) に
分割(G1: N=230, G2: N=463,
G3: N=731, G4: N=576)

	問1	問2	問3	問4	id	domain
受験者1	1	1	1	0	問1	数と計算
受験者2	1	na	1	0	問2	量と測定
受験者3	1	0	1	na	問3	図形
受験者4	1	0	0	na	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮

アソシエーションルール (AR)

「Aの正解とBの正解が共起しやすい」

前処理：

アイテム化 *_r(正誤) / *_na(無回答)

フィルタ：

最小支持度：相対 0.05 (絶対20件)

フィルタ： $|\phi| \geq 0.14$ or $|\log Lift| \geq 0.14$

有意性：Fisher片側検定 → **BH-FDR**

可視化：ラベル= ϕ (色：正=赤/負=青)

因果 (DirectLiNGAM)

「Aを正解したらBも正解する？」

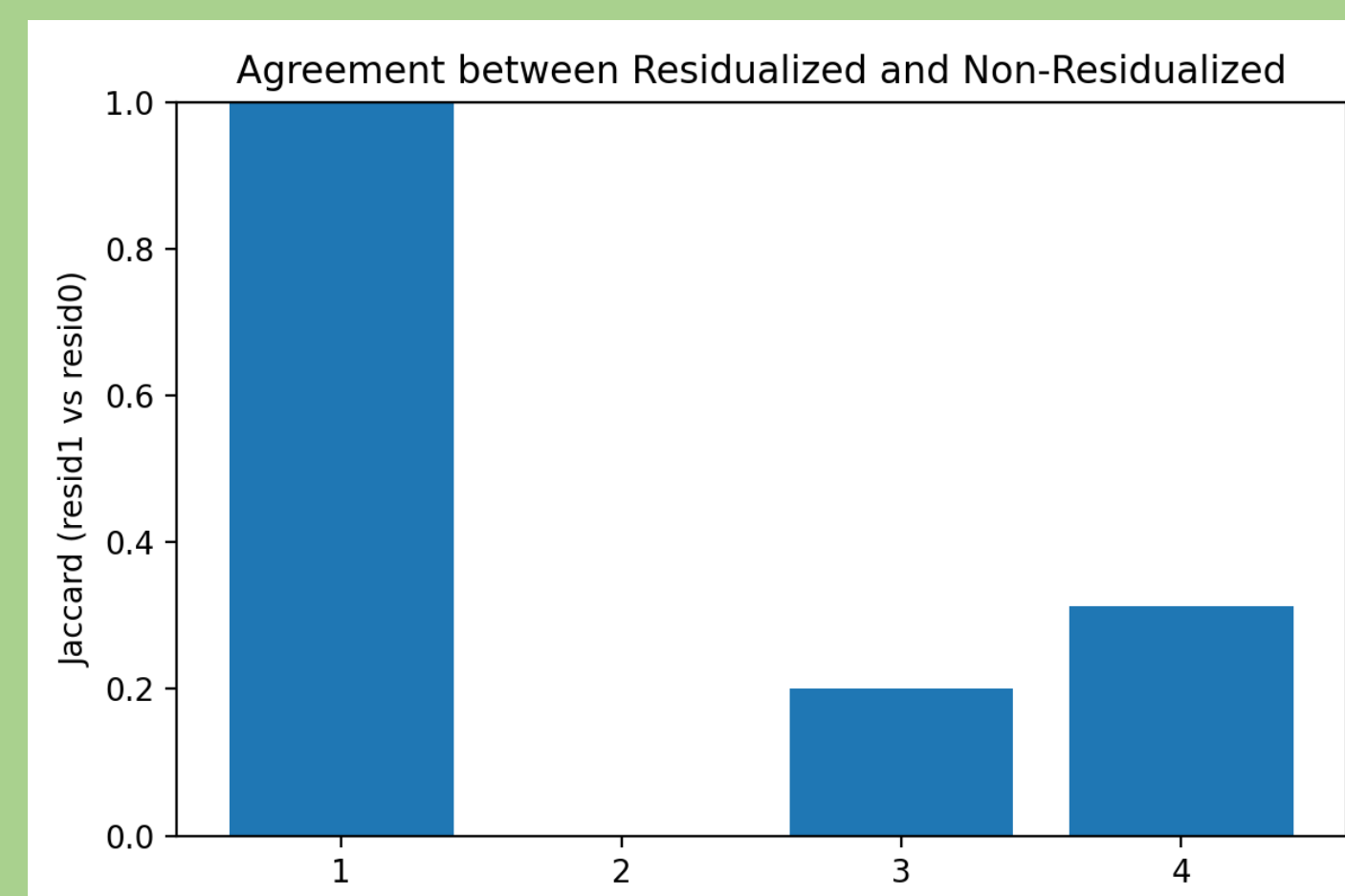
前処理：能力残差化(LOO)で一般能力を制
御。既定ではNAを設計行列に含めない。
ブートストラップCI(n=600, 95%CI：0
を跨がない)

効果量 $\Delta p = \beta \times (\sigma Y / \sigma X)$ 、 $\Delta p \geq 0.08$

可視化：ラベル= Δp (色：正=赤/負=青)

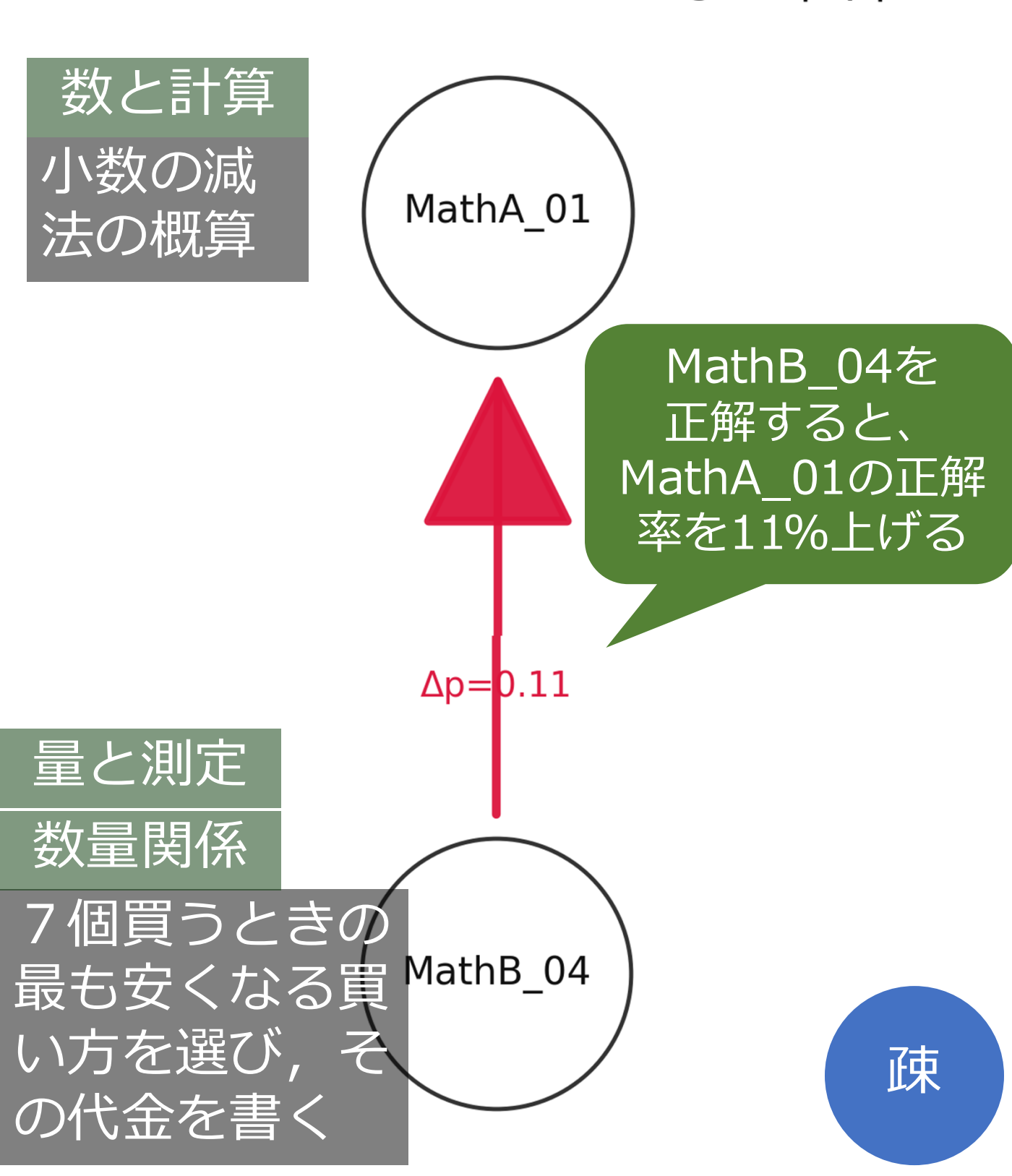
出力

- HTML + JSON (ノード座標・エッジ指標) → PNG。
- 群内比較：残差あり vs なし
(重なり/Jaccard・方向反転)
- 群間比較：G1↔G4 などの差異
(ドメイン別も)
- ARules vs LiNGAM の対応関
係も点検。

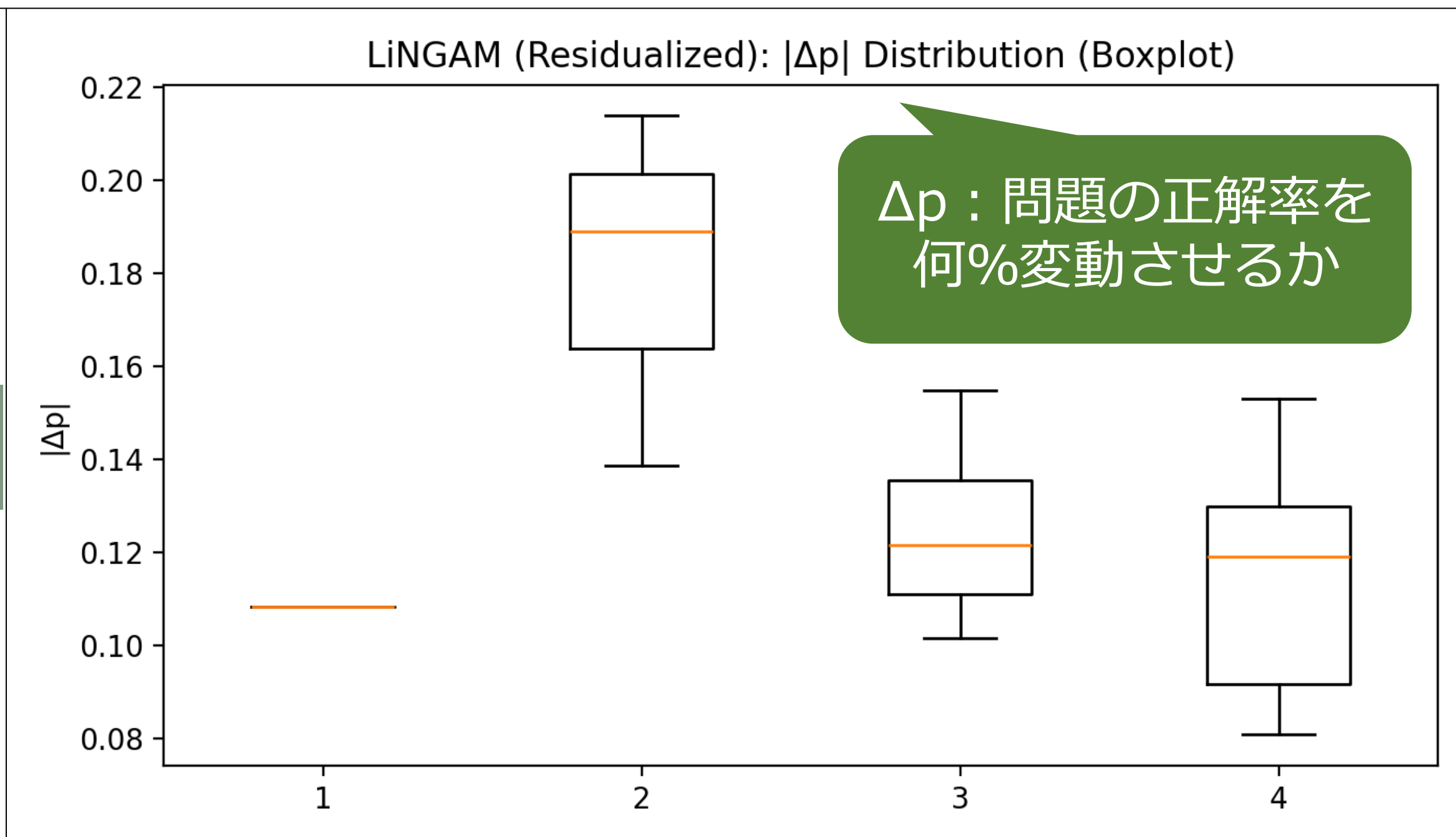
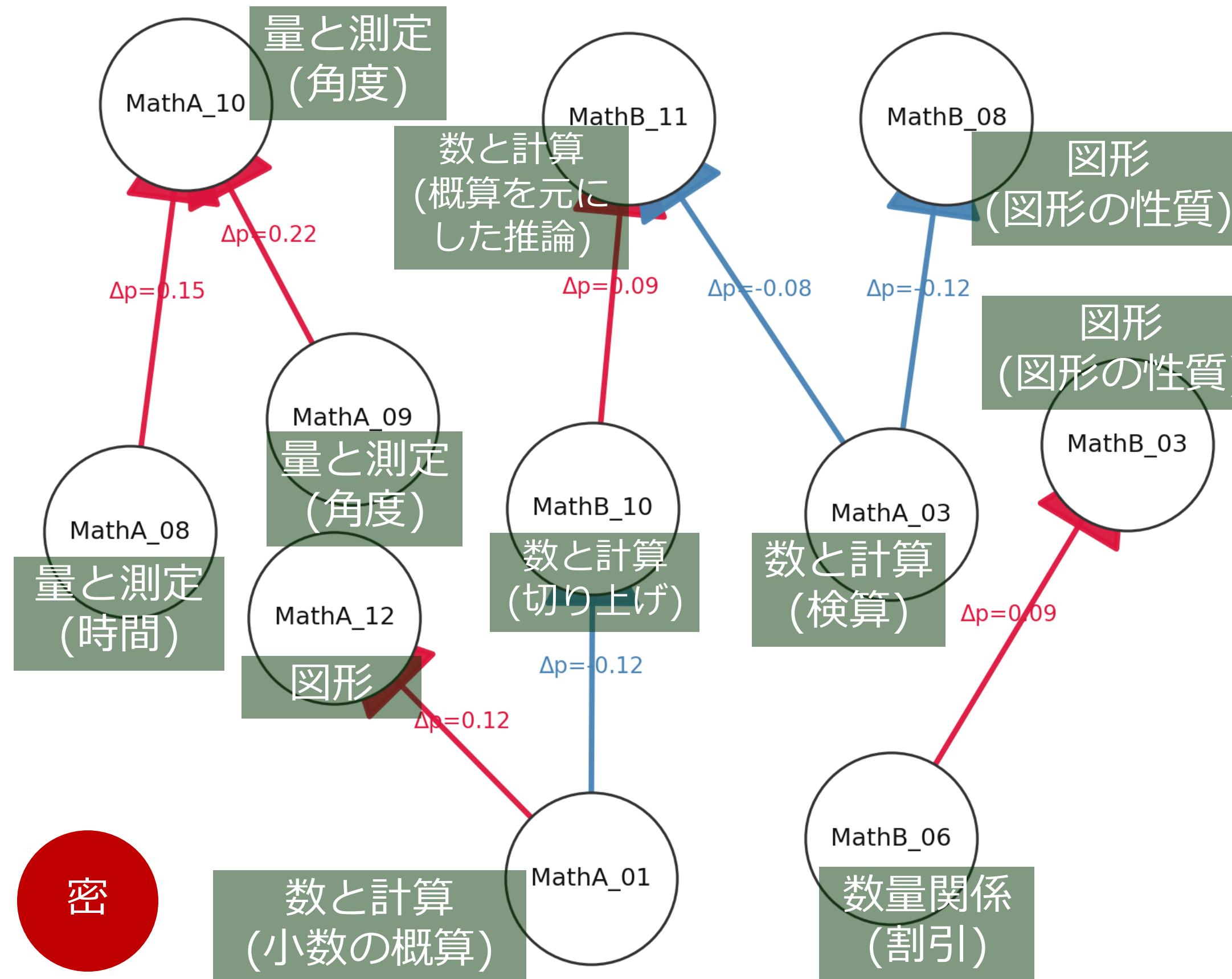


結果と考察

Group 1 (DirectLiNGAM, LOO residual)
Causal (residualized) — 1 edges ($|\Delta p| \geq 0.0$)



Group 4 (DirectLiNGAM, LOO residual) — Causal (residualized) — 8 edges ($|\Delta p| \geq 0.0$)



- 効果量 ($|\Delta p|$) の分布は群によって異なる
G2の $|\Delta p|$ が最も大きく、G3/G4は中程度、G1は小さい
1点のみ。→ **中位群(G2)は設問間の影響が相対的に強
く、指導介入の余地が大きい可能性。**

- 高学力群(G1)は疎、低学力群(G4)は密** 能力残差化後のLiNGAMで、G1は**有意エッジ1本**のみ、G4は**8本**の有意エッジを検出。
→ 能力を取り除くと、高学力群では問題間の直接的な因果関係が残らず、**低学力群ではつまずきが連鎖しやすい**構造が顕在化。
- 負のエッジは主にG4で出現** G4では負の Δp (青)が複数観測され、**ある問題の正解が別問題の正解確率を下げる**関係が存在。
→ **操作手続きの衝突・方略の誤解**などの学習上の干渉、時間配分などの取り組み方のミスを示唆。

限界と今後

- 観察データの限界**：得られたネットワークはあくまで推定。現場での介入等の検証が必要。
- Δp 閾値・残差化の仕様**：閾値を下げることで、効果量は低いが有意な関係性を探索的に探せる。より適切な残差化手法を検討可能(本研究では線形回帰を使用。ロジスティック回帰や他の手法も試すことで、より実態に則した残差化にできる可能性)。
- 一般化**：他教科・学年・他年度での再現性検証。→カリキュラム設計や学年間の接続を改善。教育施策のエビデンスを強化。