

# 高校数学の知識ネットワークによる単元構造の可視化とその応用

## Knowledge Network of High School Mathematics and Its Application

谷口 陽聖†  
Yosei Taniguchi

### 1. 緒言

近年情報化が進み、知識を関連付けて学習することが重要視されている。しかし、どのように知識を関連付けるかが不明瞭である。また我が国は中学生の数学的リテラシーにおいて世界トップレベルであるにも関わらず、高校以降の数学に対する苦手意識が増加している<sup>[1]</sup>。そのため、高校数学の単元のつながりを知識のつながりとみなした知識ネットワークを解析することで、理解度の高い学習者はどのように知識に関連をもたせて学習しているか明らかにする必要がある。しかし、この知識ネットワークは主観的に構築されることから、理解度の高い学習者であっても異なる知識ネットワークを構築する。また単元間のつながりは直接的なつながりと、複数の単元を経由してつながる間接的なつながりが存在する。さらにつながりごとに重要度が異なるため、単元間のつながりの中からより効果的に重要度の高いものを選び出す作業が困難となっている。

本研究では、この知識ネットワークにグラフ理論を用いた構造モデリング手法を応用した。また、高校数学の理解度が高い人(習熟者)と理解度が低い人(非習熟者)の分類を行うため、高校数学の教科書の基本レベルの学力試験を実施することで、学力試験結果を用いた習熟者・非習熟者の分類をし、解析を行った。習熟者・非習熟者間の特徴の比較と単元構造の可視化を用いて、学習者の支援を目指す。

### 2. 知識ネットワーク解析手法

#### 2.1 解析対象

高校数学の指導者 10 名と、ある大学の新生 55 名の知識ネットワークを解析した。本研究では文部科学省の学習指導要領に基づいた 24 単元をノード、それらの単元同士が関連するかをリンクとした知識ネットワークを構築した。

#### 2.2 データの解析手法

本研究では、ネットワークのリンクの有無を表現した隣接行列を用いる。隣接行列はノード  $v_i$  からノード  $v_j$  にリンクがあるときに  $i$  行  $j$  列に 1、リンクがないときに 0 とする。このときリンクの強さを考慮し、1 を入力する部分に重みの数値を入力することで直接影響行列を作成できる。隣接行列(または直接影響行列)を全成分の和で除したものを正規化直接影響行列  $X$  と呼ぶ。ノード  $v_i$  から  $n$  個のノードを経由してノード  $v_j$  に影響する間接影響の強さは、行列  $X$  の  $n+1$  乗の  $ij$  成分で表される。これより、直接的影響と間接的影響をともに含めた総合影響行列  $T$  は

$$T = X + X^2 + X^3 + \dots = X \cdot (I - X)^{-1} \quad (1)$$

となる。ここで  $I$  を単位行列とする<sup>[2]</sup>。

#### 2.3 単元特性値

総合影響行列  $T$  を用いて中心度、原因度といったノードの特性値を算出することができる。この特性値は、そのノ

ードが構造の中でどの程度中心的な役割を果たしているか、ほかのノードに対して原因要因として作用しているか、結果要因として作用しているかの指標となる<sup>[2]</sup>。

#### 2.4 コミュニティ抽出手法

本研究ではグラフ構造からコミュニティを抽出する方法としてモジュラリティという指標を用いた。モジュラリティは、切り出した部分グラフがランダムグラフと異なる程度を示す指標であり、切り出した部分グラフのモジュラリティが高ければ、適切にコミュニティを抽出していることを示す。モジュラリティ  $Q$  は次式で定義する。

$$Q = \sum_{i \in S} \left\{ \frac{e_{ii}}{2m} - \left( \sum_{j \in S} \frac{e_{ij}}{2m} \right)^2 \right\} \quad (2)$$

ここで、抽出したコミュニティ集合を  $S$ 、コミュニティ  $i$  からコミュニティ  $j$  へ接続されているリンク数を  $e_{ij}$ 、グラフ構造全体に含まれるリンクの総数を  $m$  とする。モジュラリティの値が最大となるコミュニティ分類を発見することでコミュニティ抽出を行う<sup>[3]</sup>。本研究では重み付き有向グラフである総合影響行列に対して行うため、 $e_{ij}$  をコミュニティ  $i$  からコミュニティ  $j$  へ接続されているリンクの重みの和とし、 $m$  をグラフ構造全体に含まれるリンクの重みの総和とすることでモジュラリティを算出する。

#### 2.5 習熟者・非習熟者分類方法

習熟者と非習熟者の分類を行うため、高校数学の教科書の基本レベルの学力試験のある大学の新生 55 名に行った。試験は全 24 問で、それぞれの問題は本研究で用いた 24 単元に対応する。知識ネットワークの違いを見るため、習熟者と非習熟者の分類基準には、学力試験の平均値  $\mu$  と標準偏差  $\sigma$  を用いて、分類の基準点を  $\mu$ 、 $\mu + \sigma$  の 2 種の基準を用いてそれぞれで解析を行った。

### 3. 結果と考察

#### 3.1 データ収集手法と解析手法の妥当性

高校数学の指導者 10 名に 2.2 節の隣接行列の入力を行ってもらい、総合影響行列に変換した。さらに 2.3 節で述べた単元特性値を算出し、自身の単元特性値を確認してもらった。その結果 10 名中 10 名が自身のイメージする単元の重要度を反映できていると述べたアンケート結果を得た。ここから総合影響行列が表現する単元構造は習熟者がイメージする単元の関係を忠実に再現すると考えられる。したがって、本実験で用いたアンケート方法の有効性、総合影響行列を用いた解析の有効性が示せたといえる。

習熟者と非習熟者の各単元間リンクの違いについて解析するため、隣接行列の各単元間リンクの選択率、総合影響行列の各単元間リンクの平均値に有意な差があるか検定を行った。習熟者と非習熟者の分類の基準点を  $\mu$  と  $\mu + \sigma$  で行ったところ解析結果は異なる部分が多く、特徴的な単元間リンクを見出すことはできなかった。そのため単元構造の可視化を行うことで、知識ネットワーク全体の解析を行う。

†同志社大学, Doshisha University

### 3.2 可視化とコミュニティ抽出解析

可視化することで、知識ネットワークの全体像を把握することができる。分類基準点を $\mu$ としたときの、習熟者と非習熟者の総合影響行列を可視化した図をそれぞれ Fig.1, Fig.2 に示す。ここでノードの大きさは中心度に、リンクの大きさは重みの大きさに対応する。また 2.4 節の手法で得たコミュニティ毎にノード及びリンクの色を変えた。さらに見やすさのために、重みの小さいリンクは意図的に非表示としている。

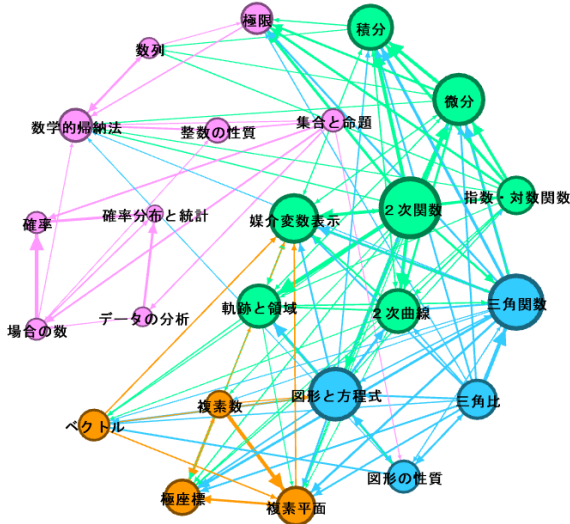


Fig.1 A visualization of experts' knowledge network

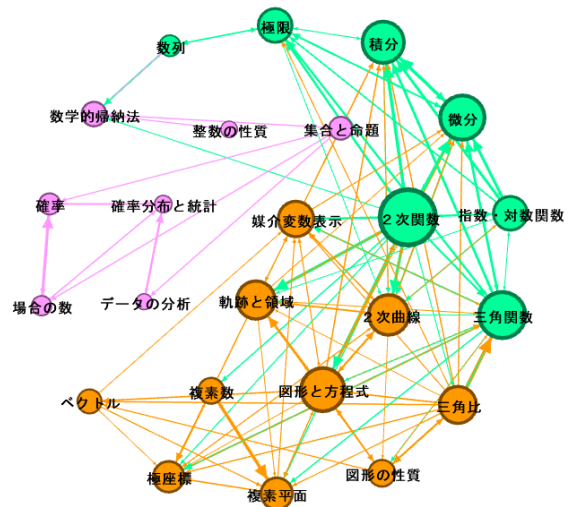


Fig.2 A visualization of non-experts' knowledge network

また習熟者と非習熟者のコミュニティ抽出結果をそれぞれ Tab.1, Tab.2 に示す。C<sub>1</sub> はコミュニティの分類番号である。

Community	Topics
C <sub>1</sub>	2次関数, 軌跡と領域, 指数・対数関数, 微分, 積分, 2次曲線, 媒介変数表示
C <sub>2</sub>	三角比, 図形と方程式, 三角関数, 図形の性質
C <sub>3</sub>	複素数, 複素平面, 極座標, ベクトル
C <sub>4</sub>	集合と命題, データの分析, 極限, 場合の数, 確率, 整数の性質, 数列, 数学的帰納法, 確率分布と統計

Tab.2 Result of clustering with non-experts' knowledge network

Community	Topics
C <sub>1</sub>	2次関数, 三角関数, 指数・対数関数, 微分, 積分, 極限, 数列
C <sub>2</sub>	三角比, 複素数, 図形と方程式, 軌跡と領域, 2次曲線, 複素平面, 媒介変数表示, 極座標, 図形の性質, ベクトル
C <sub>3</sub>	集合と命題, データの分析, 場合の数, 確率, 整数の性質, 数学的帰納法, 確率分布と統計

Tab.1 より、習熟者の C<sub>1</sub> は連続関数の解析を行う単元、C<sub>2</sub> は基本的な手法で図形を扱う単元、C<sub>3</sub> は応用的な手法を用いて図形を扱う単元、C<sub>4</sub> は論理的に推論を進めるための学習に役立つ単元が集まる傾向があるといえる。これらのコミュニティは従来の分野分けである、「数学 I」「数学 II」「数学 III」「数学 A」「数学 B」が入り混じっており、新しい単元分類の可能性もある。また、習熟者のコミュニティ数は4であるのに対し、Tab.2 より非習熟者のコミュニティ数は3となっている。これは Tab.1 の C<sub>2</sub> と C<sub>3</sub> の単元が同一のコミュニティに統合されて Tab.2 の C<sub>2</sub> に分類されている傾向がある。Tab.1 の C<sub>2</sub> と C<sub>3</sub> はともに図形を扱う単元のコミュニティであるため、統合されやすいことがわかる。この他にもコミュニティ抽出の違いからそれぞれのグループの特徴を見つけ出すことができる。

本来知識ネットワークには全ての単元間にリンクがあるため、Fig.1 や Fig.2 のような知識ネットワークから重要とされるリンクを一目で判断することが困難となっている。そのため本研究で得た習熟者の知識ネットワークを参考にした重要度順に単元間リンクをたどり、理解する学習方法を提案する。この方法では、ある重み以下のリンクは表示しないというフィルタを用い、フィルタの閾値を少しずつ下げることで徐々にリンクを表示する。追加される単元間リンクの関係性を考えることで、習熟者の重要度が高い単元間リンクから順に単元間の関係性を理解することができる。このように可視化を用いて、学習者が知識ネットワークを構築する支援ができると考える。

### 4. 結言

本稿では、グラフ理論を用いた知識ネットワークの解析手法の妥当性を示した。またコミュニティ抽出解析を行い、Gephi による可視化を用いて、知識ネットワークを利用した学習方法の提案を行った。今後は、提案した学習方法を実際に学習者に実行してもらい、その有効性を検討していく。また、Knowledge Tracing を用いることで、学力試験結果から知識のつながりを読み取り、知識ネットワークと組み合わせることを検討している。

### 参考文献

- [1] 文部科学省国立教育政策研究所, OECD 生徒の学習到達度調査 (PISA2018) のポイント, [https://www.nier.go.jp/kokusai/pisa/pdf/2018/01\\_point.pdf](https://www.nier.go.jp/kokusai/pisa/pdf/2018/01_point.pdf), (2019-12-10).
- [2] Wang, W. et al., DEMATEL-based model to improve the performance in a matrix organization, Expert Systems with Applications, Vol.39, Issue.5, (2012).
- [3] Newman, M. E. J., Girvan, M., Finding and evaluating community structure in networks, Physical Review E, Vol.69, (2004), [org/10.1103/PhysRevE.69.026113](https://doi.org/10.1103/PhysRevE.69.026113).