

## ネットワーク解析のための構成論的な参照基準モデル

高口 太朗

LINE 株式会社 Data Labs

ネットワーク解析とは、事物間の関係性を表したネットワーク（グラフ）型のデータを操作することによって適切な低次元の構造を取り出すための試みである。たとえば、ネットワークの頂点ごとに接続する枝の本数（次数とよばれる）を数えることは、ネットワークを頂点の次数を並べた1次元のベクトルに写す操作であるとみなせる。この次数のベクトルは、ネットワーク全体の枝の密度や頂点ごとの次数のばらつきについての情報をもつ。あるいは、ネットワークに対応するラプラシアン行列の固有値と固有ベクトルは、ネットワーク上の拡散についての情報をもつ低次元の表現である。ネットワーク解析は計算機科学、数学、物理学などの広範な分野に関連がある。

あるネットワーク解析の手法が有用であるか、またはあるネットワークに手法を適用した結果が有用なものであるか、どのように判断すればよいであろうか。ひとつの方針は、参照基準となるネットワークモデルを用意することである。たとえば、頂点数と枝数を固定した上でそれ以外の構造を完全にランダムとするモデルは Erdős-Rényi ランダム・グラフとよばれる。Erdős-Rényi ランダム・グラフの次数分布と実ネットワークの次数分布を比較することにより、ランダムなつながりではまず起こりえない極端な次数をもつ頂点をあぶり出すことができる。

現実的な参照基準モデルには、頂点数と枝数を固定する以上の構造をもつことが要請される。2000年前後から盛んになったネットワーク科学研究のこれまでの知見から、現実のネットワークの多くが共通の構造的特徴を示すことが明らかとなった。そのような構造的な特徴には次数のスケールフリー性、次数相関、コミュニティ構造、階層構造などがある。これらの特徴を再現するモデルを構成することにより、より現実的な比較対象を得たい。

本講演では、指数ランダム・グラフモデルと確率的ブロックモデルという2つの構成論的な参照基準モデルを中心に、ネットワーク科学研究における近年の発展と実データへの応用事例について概説する。これらのモデルを容易に取り扱えるソフトウェアと、発展的な話題として時間変化するネットワークデータに対する参照基準モデルの研究についても紹介したい。