

Analysis and quantitative evaluation of large networks using
spectral distribution

スペクトル分布を用いた大規模ネットワークの分析と定量的
評価について

by

Naoki Murakami

村上直輝

A Senior Thesis

卒業論文

Submitted to

the Department of Information Science

the Faculty of Science, the University of Tokyo

on January 29, 2021

in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Science

Thesis Supervisor: Hiroshi Imai 今井浩

Professor of Information Science

ABSTRACT

In analyzing the properties and structure of a network, the spectrum can be a very important tool. However, since computing the full eigendecomposition is expensive, the analysis of large networks have been done by using only a part of the full spectrum. Recently, a method for calculating the entire spectrum distribution has been developed, but its application has not yet been deeply considered, and it has been limited to qualitative evaluation by visualizing the spectrum distribution.

In this study, we will analyze the distribution of the overall spectrum of the theoretically created graphs and the real-world network. Then, we will discuss the relation between the properties of the network and the spectrum. As a result, we find that graphs in which vertices are adjacent to each other with similar properties have sharp spectra found in complex networks. Furthermore, we illustrate that graphs with randomness in vertex adjacencies will have the semicircular spectra found in random graphs. We also suggest applying the concepts of entropy and divergence of discrete probability distributions to quantitative evaluation and comparison of spectral distributions. A fast way to search for similar spectra using cosine distance is also discussed.

論文要旨

ネットワークの性質や構造を分析する上で、スペクトルは非常に重要な指標となり得る。しかし、大規模ネットワークについては、計算量の問題からスペクトルの一部を利用した分析しか行われていなかった。近年スペクトルの分布全体を計算する手法が考案されたが、応用方法などはまだ深く考えられておらず、スペクトル分布を視覚化して定性的な評価をするにとどまっている。

本研究では、理論的に得られるグラフと実世界のネットワークのスペクトル全体の分布を分析し、ネットワークの性質とスペクトルの対応について考察する。結果として、性質が近い頂点同士が隣接するようなグラフは、複雑ネットワークに見られるような鋭いスペクトルを持つこと分かった。さらに、頂点同士の隣接にランダム性があるグラフは、ランダムグラフに見られる半円状のスペクトルになることを例示する。また、スペクトル分布の定量的な評価や比較のために、離散確率分布に対するエントロピーやダイバージェンスの概念を流用することを提案する。これにより、ネットワークの構造に基づいた直感的な比較が可能になる。cosine distance を用いて類似のスペクトルを高速に検索する方法についても言及する。