Internship / master project topic: Maximum entropy graphons under degree distribution and clustering coefficient constraints

Jan Ramon

October 2017

1 Context

Graphons allow for representing distributions over large graphs. They also allow to study models of random graphs under constraints, see e.g. (Chatterjee & Diaconis 2013).

From a machine learning point of view, it is interesting to gain a better understanding of the distributions of various properties in random graphs satisfying specified pattern frequencies. This project considers a slightly different problem, where we are interested in the interaction between pattern frequency constraints and degree distribution constraints. Degree distributions are well-known characteristics of global network structure (Newman 2010)

2 Objective

We will study the problem where as input we get a degree distribution and the frequency of one or more subgraph patterns, and as output we should give the expected frequency of another subgraph pattern in a random graph satisfying the specified constraints.

3 More information

The project will consist of several steps:

- Study relevant literature

- Consider instantiations of the problem of increasing difficulty. E.g., first consider a powerlaw degree distribution and a clustering coefficient (triangle density) constraint, and aim at calculating the number of 4-cliques. Next, more general problems can be considered.

- For each considered setting, aim at answering the following questions:

- can we compute the desired expectation in a closed form or set up an approximation procedure?
- If the latter is possible, how efficiently can we approximate the desired expectation?
- For finite random graphs, what is the variance?

4 Requirements

A good background in mathematics and strong creativity is required.

References

(Chatterjee & Diaconis 2013) S. Chatterjee and P. Diaconis, Estimating and understanding exponential ran- dom graph models. The Annals of Statistics, 2013.

(Newman 2010) M. Newman, Networks an introduction, Oxford university press, 2010.