Spectral Thresholds in the Stochastic Block Model

Laura Florescu

March 29, 2015

We consider a bipartite stochastic block model on vertex sets V_1 and V_2 of size n_1 and n_2 respectively, with planted partitions in each. An algorithm to recover the partition of V_1 in the case $n_2 \gg n_1$ was used by Feldman, Perkins, and Vempala to solve instances of planted random k-SAT and planted hypergraph partitioning. The algorithm required $\tilde{O}(\sqrt{n_1 n_2})$ edges to recover the partition. It was left as an open question whether the straightforward spectral approach of computing the top singular vector of the centered adjacency matrix would suffice at the same edge density.

We show the answer is no: the spectral approach requires $\Omega(n_1^{1/3} n_2^{2/3})$ edges. Nevertheless, we give two modifications of the adjacency matrix that lead to simple spectral algorithms that match the previous bound of $\tilde{O}(\sqrt{n_1 n_2})$.

Finally, we locate a sharp threshold for detection of the partition, in the sense of the results of Mossel, Neeman, Sly and Massoulié for the stochastic block model. This gives the best known bounds for planted k-SAT and hypergraph partitioning as well as showing a barrier to further improvement via the reduction to the bipartite block model. Joint work with Will Perkins.