Statistically and Computationally Optimal Estimation and Inference in Common Subspaces

Joshua Agterberg

Department of Statistics, University of Illinois Urbana-Champaign

ABSTRACT

In this talk we investigate the statistical and computational limits for the common subspace model, a model wherein one observes a collection of symmetric low-rank matrices perturbed by noise, where each low-rank matrix shares the same common subspace. First, we propose an estimator based on projected gradient descent initialized via a spectral sum of squared matrices and show that it achieves the optimal sin Theta error under a strong signal-to-noise ratio (SNR) condition, and we further give evidence that this SNR condition is necessary for a polynomial-time estimator to exist. Next, we turn to estimation and inference for the sin Theta distance itself, and we show that our estimator achieves an asymptotically Gaussian distribution with a bias term that vanishes under a strong signal requirement. Based on this limiting result we propose confidence intervals and show that they are minimax optimal, though the resulting confidence intervals require knowledge of the SNR. We then turn to designing adaptive confidence intervals for the sin Theta error, and we show that adaptivity is information-theoretically impossible unless the SNR is sufficiently strong. Consequently, our results unveil a novel phenomenon: despite the SNR being "above" the computational limit for estimation, adaptive statistical inference may still be information-theoretically impossible.

Keywords: networks, spectral methods, multilayer networks

UBSea: A Unified Community Detection Framework

Xiancheng Lin[1], **Hao Chen[1]**

[1]*University of California, Davis*

ABSTRACT

Detecting communities in networks is an important task across many disciplines, including statistics, social science, and engineering. Existing approaches often target one community structure at a time -- assortative in most cases, and to a lesser extent, disassortative or core-periphery -- without offering a simple, unified framework across all cases. We propose UBSea (Unified Bigroups Standardized Edge-count Analysis), a strategic extension of modularity that unifies the detection of all three mixing patterns. UBSea automatically identifies the appropriate structure and applies to both directed and undirected networks. We establish weak and strong consistency results under the stochastic block model (SBM), expanding the class of graphs where consistent recovery can be achieved by efficient algorithms, including asymmetric and core-periphery settings not covered by previous guarantees. Extensive simulations under the SBM and degree-corrected SBM confirm UBSea's high accuracy across diverse scenarios, and applications to benchmark datasets demonstrate that while other methods can perform well in individual regimes, UBSea provides a single framework that performs reliably across assortative, disassortative, and core–periphery networks.

Keywords: Assortative mixing; Disassortative mixing; Core-periphery structure; Consistency and asymptotics

Finding Anomalous Cliques in Inhomogenous Networks using Egonets

Subhankar Bhadra[1], **Srijan Sengupta[2]**

[1]*Pennsylvania State University*

[2]*North Carolina State University*

ABSTRACT

We consider the problem of finding an anomalous clique, i.e., a fully connected subgraph, hidden in a large network. There are two parts to this problem: (1) detection, i.e., determining whether an anomalous clique is present, and (2) identification or localization, i.e., given that an anomalous clique is detected in part 1, determining which vertices of the network constitute the clique. This problem has a number of practical applications, such as financial trading networks, brain networks, and online social networks. A rich literature already exists on the detection problem when restricted to homogeneous Erd˝os–R´enyi random graphs. However, currently, no method exists that can solve the detection and identification/localization problems for inhomogeneous networks in finite time. We propose an inferential tool based on egonets to address this gap. The proposed method is computationally efficient and naturally amenable to parallel computing and easily extends to a wide variety of inhomogeneous network models. We establish the theoretical properties of the proposed method and demonstrate its empirical performance through simulation studies.

Keywords: Random graphs; Clique; Motif detection; Hypothesis testing

Back to Sessions List