Community Detection with Common Structure*

Kiyeon Jeon
kiyeonj@utexas.edu
Department of Electrical and Computer Engineering
The University of Texas at Austin

1 Introduction

The stochastic block model(SBM) is a useful method for community detection. Solving the maximum likelihood estimator(MLE) of the SBM is NP-hard, so we need convex relaxation to solve the MLE problem. The SDP relaxation of the MLE is a tractable problem, and it succeeds in exact recovery even with weak assortativity condition [1]. Also, the consistency of the SDP problem in a sparse network is proved by Grothendieck's inequality [2, 3], and it is solved by the interior point method in polynomial time [4].

In this paper, we focused on multiple networks. Even if detecting a community in a single network is meaningful, extracting communities in multiple networks is crucial in broad networks. In order to cluster multiple networks, spectral method and graph moment method were proposed depending on the existence of node correspondence [5]. However, it has not been studied yet to use a shared structure in multiple networks where multiple networks have a node correspondence. We propose a new formulation to extract a common structure between multiple networks.

2 Problem Setting

2.1 Multiple Networks

Let multiple adjacency matrices be $A_1, \ldots, A_N$, and the corresponding membership matrices be $Z_1, \ldots, Z_N$. For a simple model, we assume that the symmetric probability matrices of all networks are the same. That is to say, $A_i \sim \text{Bernoulli}(Z_i \Psi Z_i^T)$ for $i \in [N]$. When we have multiple networks, we can solve each network by SDP-3 relaxation. [1, 3]:

$$\text{for each } i,\ \hat{X}_i := \arg \max_{\substack{\text{diag}(X_i)=1_n \\ 0 \le X_i \le 1,\, X_i \succeq 0}} \langle A_i, X_i \rangle - \lambda_i \langle E_n, X_i \rangle$$

Our model is to deal with multiple networks where each network came from various sources with the same node correspondence. The multiple networks problem is formulated by imposing a penalty term on the common parts of the networks. Let our cluster matrix X_i be $C_i + Y_i$ where C_i is common structure. So, we can solve multiple networks problem by making these common structures similar:

$$\hat{C}_i, \hat{Y}_i := \arg\max \sum_{i=1}^{N} \langle A_i - \lambda_i E_n, C_i + Y_i \rangle - \sum_{i \neq j} \lambda_{ij} \|C_i - C_j\|_F^2$$

*This is course project for 2018 Fall UT-Austin graduate course SDS 385 Stat Models for Big Data

However, it is hard to assign weight parameters as a distance metric of common structures. Also, it is essential to make common sparsity structures similar. By leveraging a block sparsity that utilizes shared sparsity structure on C_i's [6], we can solve the multiple network problem efficiently. We use a method of fast projection onto a mixed-norm ball for computational efficiency [7]. And, the condition $C_i + Y_i \in \{0,1\}$ and $C_i \circ Y_i = 0$ is relaxed to $C_i, Y_i \geq 0$ and $C_i + Y_i \leq 1$.

$$\hat{C}_i, \hat{Y}_i := \arg \max_{\substack{\mathrm{diag}(C_i+Y_i)=1_n \\ C_i+Y_i\leq 1,\, C_i+Y_i\succeq 0 \\ C_i,Y_i\geq 0,\, \|C\|_{1,\infty}\leq R}} \sum_{i=1}^{N} \langle A_i - \lambda_i E_n, C_i + Y_i \rangle$$

where $C = [vec(C_1)|\cdots|vec(C_N)] \in \mathbb{R}^{n^2\times N}$ and $\|C\|_{1,\infty} = \sum_{i=1}^{n^2} \|C^i\|_\infty$ C^i is $i-$th row of C.

From the above formulation, we get $\hat{C}_i, \hat{Y}_i$ alternatively as follows:

$$\begin{aligned}
\hat{C}_i &:= \arg \max_{\substack{\mathrm{diag}(C_i)\leq 1_n \\ 0\leq C_i\leq 1, C_i\succeq 0 \\ \|C\|_{1,\infty}\leq R}} \sum_{i=1}^{N} \langle A_i - \lambda_i E_n, C_i \rangle \\
\hat{Y}_i &:= \arg \max_{\substack{\mathrm{diag}(\hat{C}_i+Y_i)=1_n \\ 0\leq \hat{C}_i+Y_i\leq 1,\, \hat{C}_i+Y_i\succeq 0}} \langle A_i - \lambda_i E_n, Y_i \rangle
\end{aligned}$$

We will solve the above problem with two steps: (1) solve $\hat{C}_i = \arg\max_{C_i\in\mathcal{F}_i} \langle A_i, C_i \rangle$ and (2) projection $C = [vec(C_1)|\cdots|vec(C_N)]$ onto $\|\cdot\|_{1,\infty} \leq K$ norm ball. In summary,

$$\begin{aligned}
\tilde{C}_i &:= \arg \max_{\substack{\mathrm{diag}(C_i)\leq 1_n \\ 0\leq C_i\leq 1, C_i\succeq 0}} \langle A_i - \lambda_i E_n, C_i \rangle \quad \text{for each } i \in [n] \\
\hat{C} &:= \Pi_{\|\cdot\|_{1,\infty}\leq R}(\tilde{C}) \\
\hat{Y}_i &:= \arg \max_{\substack{\mathrm{diag}(\hat{C}_i+Y_i)=1_n \\ 0\leq \hat{C}_i+Y_i\leq 1, \hat{C}_i+Y_i\succeq 0}} \langle A_i - \lambda_i E_n, Y_i \rangle
\end{aligned}$$

3 Implement Results

3.1 Toy example(strongly assortative model)

We simulated with the following edge probability matrix:

$$\begin{bmatrix} 0.300 & 0.200 & 0.150 \\ 0.200 & 0.300 & 0.100 \\ 0.150 & 0.100 & 0.300 \end{bmatrix}$$

We use two blocks of sizes $\bar{n_1} = (200, 100, 100)$, $\bar{n_2} = (10, 20, 10)$ and $\bar{n_2} = (100, 100, 200)$

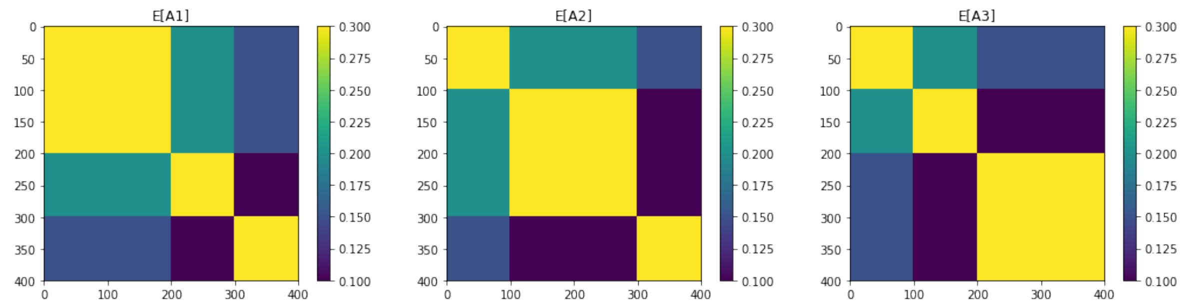

Figure 1: Population mean matrices

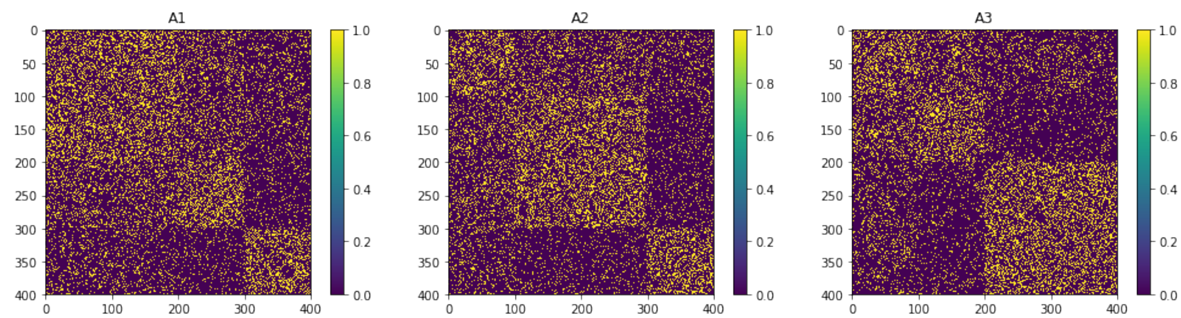

Figure 2: Adjacency matrices from SBMs

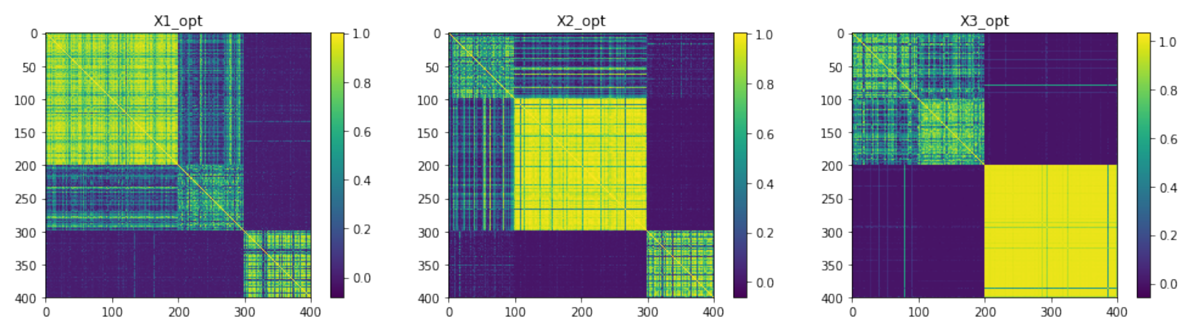

Figure 3: Results of $\tilde{C}_i$ by ADMM

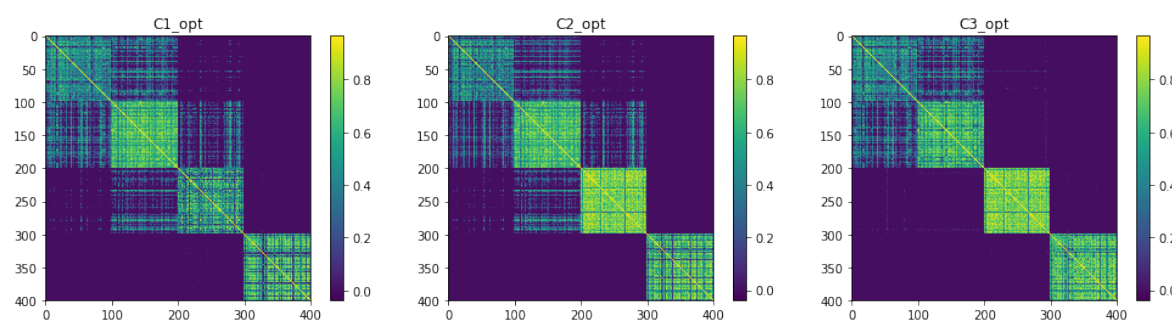

Figure 4: Results of $\hat{C}_i$ by projecting onto $1, \infty$-norm ball

References

[1] Arash A Amini and Elizaveta Levina. On semidefinite relaxations for the block model. *arXiv preprint arXiv:1406.5647*, 2014.

[2] Andrea Montanari and Subhabrata Sen. Semidefinite programs on sparse random graphs and their application to community detection. *arXiv preprint arXiv:1504.05910*, 2015.

[3] Olivier Guédon and Roman Vershynin. Community detection in sparse networks via grothendieckâĂŹs inequality. *Probability Theory and Related Fields*, 165(3-4):1025–1049, 2016.

[4] Yurii Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

[5] Soumendu Sundar Mukherjee, Purnamrita Sarkar, and Lizhen Lin. On clustering network-valued data. In *Advances in neural information processing systems*, pages 7071–7081, 2017.

[6] Sahand Negahban and Martin J Wainwright. Joint support recovery under high-dimensional scaling: Benefits and perils of âĎŞ 1,âĹđ-regularization. In *Proceedings of the 21st International Conference on Neural Information Processing Systems*, pages 1161–1168. Curran Associates Inc., 2008.

[7] Suvrit Sra. Fast projections onto mixed-norm balls with applications. *Data Mining and Knowledge Discovery*, 25(2):358–377, 2012.