Connecting the Popularity Adjusted Block Model to the Generalized Random Dot Product Graph

SDSS 2021 Lightning Presentation

John Koo

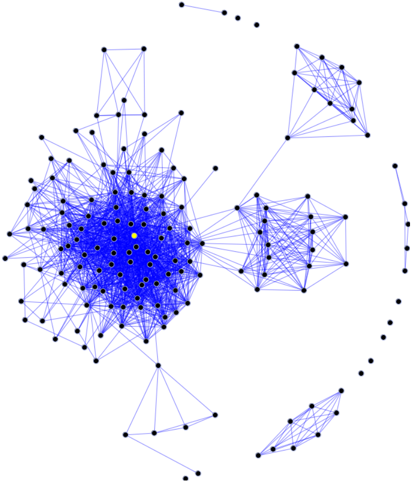
June 2021

Overview

1. Block Models and the Popularity Adjusted Block Model
2. Generalized Random Dot Product Graphs
3. Connecting the PABM to the GRDPG
4. Community Detection for the PABM

Block Models

Networks

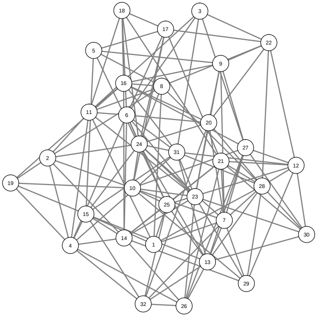


By Screenshot taken by User:DarwinPeacock - Screenshot of free software GUESS, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=6057981

Bernoulli Graphs

Let $G = (V, E)$ be an undirected and unweighted graph with $|V| = n$.
G is described by adjacency matrix A such that $A_{ij} = \begin{cases} 1 & \exists \text{ edge between } i \text{ and } j \\ 0 & \text{else} \end{cases}$
$A_{ji} = A_{ij}$ and $A_{ii} = 0 \; \forall i, j \in [n]$.
$A \sim BernoulliGraph(P)$ iff:

1. $P \in [0, 1]^{n \times n}$ describes edge probabilities between pairs of vertices.
2. $A_{ij} \overset{ind}{\sim} Bernoulli(P_{ij})$ for each $i < j$.

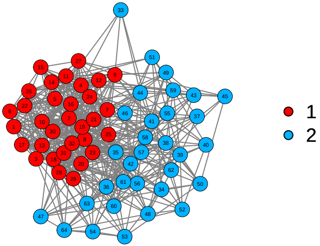


Block Models

Suppose each vertex $v_1, ..., v_n$ has hidden labels $z_1, ..., z_n \in [K]$, and each P_{ij} depends on labels z_i and z_j.
Then $A \sim BernoulliGraph(P)$ is a *block model*.

Example: Stochastic Block Model (Lorrain and White, 1971) with two communities

- $z_1, ..., z_n \in \{1, 2\}$
- $P_{ij} = \begin{cases} p & z_i = z_j = 1 \\ q & z_i = z_j = 2 \\ r & z_i \neq z_j \end{cases}$
- To make this an assortative SBM, set $pq > r^2$.
- In this example, $p = 1/2$, $q = 1/4$, and $r = 1/8$.



Popularity Adjusted Block Model

Definition based on Noroozi, Rimal, and Pensky (2020); model first proposed by Sengupta and Chen (2017).

$A \sim PABM(\{\lambda^{(kl)}\}_K)$ iff

1. w.l.o.g., organize P such that each block $P^{(kl)} \in [0,1]^{n_k \times n_l}$ contains edge probabilities between communities k and l.
2. Organize parameters as vectors such that $\lambda^{(kl)} \in \mathbb{R}^{n_k}$ are the popularity parameters of members of community k to community l.
   $\{\lambda^{(kl)}\}_K$ is the set of K^2 popularity vectors.
3. Then we can write each block of P as $P^{(kl)} = \lambda^{(kl)}(\lambda^{(lk)})^\top$.
4. Sample $A \sim BernoulliGraph(P)$.

Example: $K = 2$

$$P = \begin{bmatrix} P^{(11)} & P^{(12)} \\ P^{(21)} & P^{(22)} \end{bmatrix} = \begin{bmatrix} \lambda^{(11)}(\lambda^{(11)})^\top & \lambda^{(12)}(\lambda^{(21)})^\top \\ \lambda^{(21)}(\lambda^{(12)})^\top & \lambda^{(22)}(\lambda^{(22)})^\top \end{bmatrix}$$

Generalized Ranodm Dot Product Graphs

Generalized Random Dot Product Graph

Generalized Random Dot Product Graph $A \sim GRDPG_{p,q}(X)$ (Rubin-Delanchy, Cape, Tang, Priebe, 2020)

- Latent vectors $x_1, ..., x_n \in \mathbb{R}^{p+q}$ such that $x_i^\top I_{p,q} x_j \in [0, 1]$ and $I_{p,q} = blockdiag(I_p, -I_q)$
- $A \sim BernoulliGraph(X I_{p,q} X^\top)$ where $X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^\top$

If latent vectors $X_1, ..., X_n \overset{iid}{\sim} F$, then we write $(A, X) \sim GRDPG_{p,q}(F, n)$.

(Generalized) Random Dot Product Graph Model

Recovery/Estimation

Want to estimate X given A, or alternatively, interpoint distances, inner products, or angles.

Adjacency Spectral Embedding

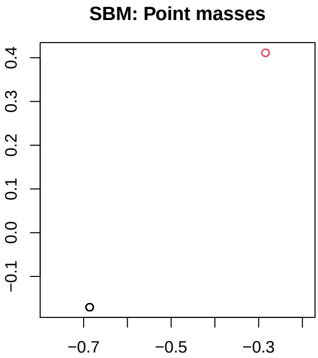
To embed in $\mathbb{R}^{p+q}$,

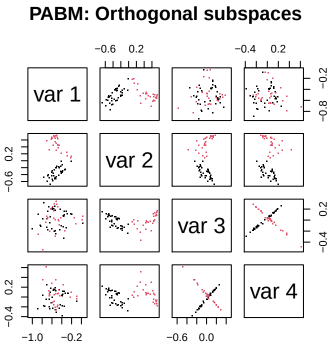
1. Compute $A \approx \hat{V}\hat{\Lambda}\hat{V}^\top$ where $\hat{\Lambda} \in \mathbb{R}^{(p+q)\times(p+q)}$ and $\hat{V} \in \mathbb{R}^{n\times(p+q)}$ by using p most positive and q most negative eigenvalues.
2. Let $\hat{X} = \hat{V}|\hat{\Lambda}|^{1/2}$.

$$\max_i \|\hat{X}_i - Q_n X_i\| \stackrel{a.s.}{\to} 0 \text{ (Rubin-Delanchy et al., 2020)}$$

$$Q_n \in \mathbb{O}(p, q)$$

Connecting Block Models to the GRDPG





Connecting the PABM to the GRDPG

Connecting the PABM to the GRDPG ($K = 2$)

Theorem (KTT): $A \sim PABM(\{\lambda^{(kl)}\}_2)$ is equivalent to $A \sim GRDPG_{3,1}(XU)$ for block diagonal X constructed from $\{\lambda^{(kl)}\}_2$ and predetermined $U \in \mathbb{O}(4)$.

Proof: Decompose P as follows

$$X = \begin{bmatrix} \lambda^{(11)} & \lambda^{(12)} & 0 & 0 \\ 0 & 0 & \lambda^{(21)} & \lambda^{(22)} \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$P = (XU) I_{3,1} (XU)^\top = \begin{bmatrix} \lambda^{(11)}(\lambda^{(11)})^\top & \lambda^{(12)}(\lambda^{(21)})^\top \\ \lambda^{(21)}(\lambda^{(12)})^\top & \lambda^{(22)}(\lambda^{(22)})^\top \end{bmatrix}$$

Connecting the PABM to the GRDPG ($K \geq 2$)

Theorem (KTT): $A \sim PABM(\{\lambda^{(kl)}\}_K)$ is equivalent to $A \sim GRDPG_{p,q}(XU)$ such that

- $p = K(K+1)/2$
- $q = K(K-1)/2$
- U is orthogonal and predetermined for each K
- X is block diagonal and composed of $\{\lambda^{(kl)}\}_K$
  $\implies$ if $x_i^\top$ and $x_j^\top$ are two rows of XU corresponding to different communities, then $x_i^\top x_j = 0$.

Remark (non-uniqueness of the latent configuration):
$A \sim GRDPG_{p,q}(XU) \implies A \sim GRDPG_{p,q}(XUQ)$
$\forall Q \in \mathbb{O}(p,q)$

Community Detection for the PABM

Sparse Subspace Clustering

Corollary: X is block diagonal by community and U is orthogonal $\implies$ each community corresponds to a subspace in $\mathbb{R}^{K^2}$.

Subspace property holds even with linear transformation $Q \in \mathbb{O}(p, q)$.

Noroozi et al. observed that the rank of P is K^2 and the columns of P belonging to each community has rank K to justify SSC for the PABM.

$$\arg\min_{c_i} \|c_i\|_1 \text{ subject to } A_{\cdot,i} = Ac_i \text{ and } c_i^{(i)} = 0$$

GRDPG-based approach: Apply SSC to the ASE of A.

$$\arg\min_{c_i} \|c_i\|_1 \text{ subject to } \hat{x}_i = \hat{X} c_i \text{ and } c_i^{(i)} = 0$$

$$A \approx \hat{X} I_{p,q} \hat{X}^\top$$

Sparse Subspace Clustering

Theorem (KTT): If $P = V \Lambda V^\top$ and $B = n V V^\top$, then $B_{ij} = 0$ $\forall i, j$ in different communities.

Theorem (KTT):

Let

- P_n describe the edge probability matrix of the PABM with n vertices, and $A_n \sim BernoulliGraph(P_n)$.
- $\hat{V}_n$ be the matrix of eigenvectors of A_n corresponding to the $K(K+1)/2$ most positive and $K(K-1)/2$ most negative eigenvalues.

Then

- $\exists N < \infty$ such that when $n > N$, $\sqrt{n} \hat{V}_n$ obeys the Subspace Detection Property with probability 1.

Orthogonal Spectral Clustering

Theorem (KTT): If $P = V\Lambda V^\top$ and $B = nVV^\top$, then $B_{ij} = 0$ $\forall i, j$ in different communities.

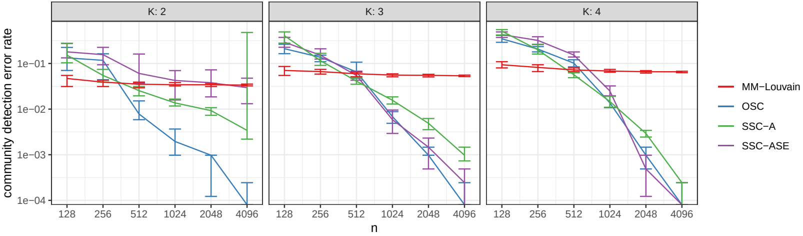
Orthogonal Spectral Clustering algorithm:

1. Let V be the eigenvectors of A corresponding to the $K(K+1)/2$ most positive and $K(K-1)/2$ most negative eigenvalues.
2. Compute $B = |nVV^\top|$ applying $|\cdot|$ entry-wise.
3. Construct graph G using B as its similarity matrix.
4. Partition G into K disconnected subgraphs.

Theorem (KTT): Let $\hat{B}_n$ with entries $\hat{B}_n^{(ij)}$ be the affinity matrix from OSC. Then $\forall$ pairs (i, j) belonging to different communities and sparsity factor satisfying $n\rho_n = \omega\{(\log n)^{4c}\}$,

$$\max_{i,j} \hat{B}^{(ij)} = O_P\Big(\frac{(\log n)^c}{\sqrt{n\rho_n}}\Big)$$

Simulation Results



IQR of community detection error rates using OSC (blue) compared against SSC on the ASE of A (purple), MM (red), and SSC on the adjacency matrix (green). Communities are approximately balanced. Simulations were repeated 50 times for each sample size.