BSTERGM: Bayesian Separable-Temporal Exponential family Graph Models

Choi Seokjun

16 Oct. 2020

Outline

Random graphs

Definition (Random graphs and related terminologies)

Let $\mathcal{B} = \{0, 1\}$. For given $n \in \mathbb{N}$,

- The set $\mathcal{Y} \subset \mathcal{B}^{n^2}$ is a set of graphs of n nodes (without weights for each nodes and edges.)
- Let Ω be an event set. We say $Y : \Omega \to \mathcal{Y}$ is a random variable for a graph, or a random graph.
- For a random graph $Y \in \mathcal{Y}$, denote the edge between i-th node and j-th node by Y_{ij} for $i, j = 1, 2, ..., n$, satisfying $Y_{ij} = 1$ if the edge is connected. Otherwise, $Y_{ij} = 0$.
- If edges of Y have directions, then Y is called a directed graph. Otherwise, Y is called a undirected graph.

Let me notate a realization of random graph by y and its edges by y_{ij} for $i, j = 1, 2, ..., n$.

Here is a remark. These are obvious that, for given $n \in \mathbb{N}$,

- $|\mathcal{Y}| = 2^{n(n-1)}$ if $\mathcal{Y}$ is the set of all directed graphs not permitting self-connecting edges.
- $|\mathcal{Y}| = 2^{n(n-1)/2}$ if $\mathcal{Y}$ is the set of all graphs of undirected one.

;thus, the size of $\mathcal{Y}$ grows exponentially when n increases.

ERGM: Exponential family Random Graphs Models

Definition (ERGM: Exponential family Random Graphs Models)

Let $\mathcal{Y}$ be a set of graphs with n nodes and Y be a random graph (on $\mathcal{Y}$.) Set a distribution on $\mathcal{Y}$ to

$$P(Y = y; \theta) = \frac{exp(\theta^T s(y))}{c(\theta)}$$

for some $\theta \in \mathbb{R}^p$, where $s(y) \in \mathbb{R}^p$ is a vector which is part of y's sufficient network statistics, and $c(\theta) \in \mathbb{R}$ is a normalizing constant satisfying $c(\theta) = \sum_{y \in \mathcal{Y}} exp(\theta^T s(y))$.
Models which have a such form called the ERGM.

Definition (TERGM: Temporal Exponential family Random Graphs Models)

Let $\mathcal{Y}$ be a set of graphs with n nodes. Let $Y_1 = y_1 \in \mathcal{Y}$ be given and $Y_2, ..., Y_T (T \in \mathbb{N})$ be random graphs (on $\mathcal{Y}$). Set a distribution on $\mathcal{Y} \times ... \times \mathcal{Y}$ ($T - 1$ folds) to

$$P(Y_t = y_t | Y_{t-1} = y_{t-1}; \theta) = \frac{exp(\theta^T s(y_t, y_{t-1}))}{c(\theta, y_{t-1})}$$

for $\theta \in \mathbb{R}^p$ and with the first-order Markov assumption $P(Y_2, Y_3, ..., Y_T | Y_1) = P(Y_2|Y_1)P(Y_3|Y_2)...P(Y_T|Y_{T-1})$, where $s(y_t, y_{t-1}) \in \mathbb{R}^n$ is a part of sufficient network statistics, and $c(\theta, y_{t-1}) = \sum_{y \in \mathcal{Y}} exp(\theta^T s(y, y_{t-1}))$ is a normalizing constant. Models which have a such form called the TERGM.

STERGM: Separable-Temporal Exponential family Random Graphs Models

Definition (STERGM: Separable-Temporal Exponential family Random Graphs Models)

Let $\mathcal{Y}$ be a set of graphs with n nodes. Let $Y_1 = y_1 \in \mathcal{Y}$ be given and $Y_2, ..., Y_T (T \in \mathbb{N})$ be random graphs (on $\mathcal{Y}$). Set a distribution on $\mathcal{Y} \times ... \times \mathcal{Y}$ ($T-1$ folds) by following way:
Let $\mathcal{Y}^+|_t$ be a subset of $\mathcal{Y}$ consisting all graphs which have equal or additional edges comparing to y_{t-1}.
Likewise, let $\mathcal{Y}^-|_t$ be a subset of $\mathcal{Y}$ consisting all graphs which have equal or sparse edges comparing to y_{t-1}.
Next, to $y_t^+ \in \mathcal{Y}^+|_t$ and $y_t^- \in \mathcal{Y}^-|_t$, set

$$P(Y_t^+ = y_t^+ | Y_{t-1} = y_{t-1}; \theta^+) = \frac{exp((\theta^+)^T s(y_t^+, y_{t-1}))}{c(\theta^+, y_{t-1})}, P(Y_t^- = y_t^- | Y_{t-1} = y_{t-1}; \theta^-) = \frac{exp((\theta^-)^T s(y_t^-, y_{t-1}))}{c(\theta^-, y_{t-1})}$$

for some $\theta^+, \theta^- \in \mathbb{R}^p$, $s(y_t^+, y_{t-1}), s(y_t^-, y_{t-1}) \in \mathbb{R}^n$, which are parts of sufficient network statistics, and normalizers $c(\theta^+, y_{t-1}) = \sum_{y^+ \in \mathcal{Y}^+} exp((\theta^+)^T s(y^+, y_{t-1})), c(\theta^-, y_{t-1}) = \sum_{y^- \in \mathcal{Y}^-} exp((\theta^-)^T s(y^-, y_{t-1}))$.
Then, defining operations $+, -$ on $\mathcal{Y}$ following the boolean algebra edgewise-ly, set y_t to

$$y_t = y_t^+ - (y_{t-1} - y_t^-) = y_t^- + (y_t^+ - y_{t-1})$$

Additionally, assume that

- The first-order Markov assumption: $P(Y_2, ..., Y_T | Y_1) = P(Y_2|Y_1)...P(Y_T|Y_{T-1})$
- The separability: the conditional independence between Y_t^+ and Y_t^- for all $t = 2, ..., T$; thus, $P(Y_t = y_t | Y_{t-1} = y_{t-1}; \theta^+, \theta^-) = P(Y_t^+ = y_t^+ | Y_{t-1} = y_{t-1}; \theta^+) P(Y_t^- = y_t^- | Y_{t-1} = y_{t-1}; \theta^-)$

Models which have a such form called the SERGM.