Social Networks Properties Study of Latent Class models

–

October 6, 2016

1 Introduction

2 MMSB

In the following, we will consider graph $G = (V, E)$ of size $|V| = N$. Hence G defines a adjacency binary matrix $Y = (y_{ij})_{i,j \in \mathbb{N}}$ such that $y_{ij} = 1$ if there an edge between node i and j. Additionally and whitout loss of generality we consider a undirected graph and it imply that $\mathbf{p}(z_{i \to j}) = \mathbf{p}(z_{j \leftarrow i})$ and $\phi_{kk'} = \phi_{k'k}$

...generative models...

Basically we have the following independences:

- $(y_{ij} \perp\!\!\!\perp Y^{-ij} \mid \phi_{kk'})$, links interactions are iid given the classes membership and weight interactions.
- $(z_{i \to j} \perp\!\!\!\perp Y^{-ij} \mid f_i)$, Each interactions draws it's own membership.
- $(\phi_{kk'} \perp\!\!\!\perp \mathbf{\Phi}^{-kk'} | \sigma_w)$, class weight interactions are iid (Beta).
- $(f_i \perp\!\!\!\perp F^{-i} | \alpha)$, feature vectors are iid (Dirichlet) over nodes of the network.

Notation remark 1: we write equivalently $\mathbf{p}(z_{i \to j} = k | f_i)$ as $\mathbf{p}(k \mid f_i)$ as the probability that node i takes membership in class k for the interactions with j.

Notation remark 2: $p(d_i^{N,l}) = p((y_{ij})_{j=0}^{N})$

2.1 Likelihoods

The likelihood that an edge is being generated is given by the observation plate in the graphical model. Marginalizing over the class assignment for both vertices gives us:

$$p(y_{ij} \mid f_i, f_j, \Phi) = \sum_{k<k'} p(y_{ij} \mid \phi_{kk'})\ p(k \mid f_i)\ p(k' \mid f_j) \tag{2.1}$$

Let's now tackle the probability of the degree d_i of a random vertex i, that is equal to n. Let's consider the joint probability from the graphical model:

$$p(d_i, F, \Phi \mid \alpha, \lambda) = p(f_i|\alpha) p(\Phi|\lambda) \prod_{j \neq i} p(f_j|\alpha) \sum_{k<k'} p(y_{ij} \mid \phi_{kk'})\ p(k \mid f_i)\ p(k' \mid f_j) \tag{2.2}$$

The degree distribution if then obtained by marginalizing the models parameters:

$$p(d_i \mid \alpha, \lambda) = \int_F \int_\Phi p(f_i|\alpha)p(\Phi|\lambda) \prod_{j \neq i} p(f_j|\alpha)p(y_{ij} \mid f_i, f_j, \Phi) \; dF d\Phi \tag{2.3}$$

As for LDA, there is no closed form expression for this integral due to complex coupling between the weights interactions $\phi_{kk'}$ and the class membership $z_{i\rightarrow \cdot}$.

2.2 Predictive Distribution

Even though the degree distribution has no closed form, the predictive likelihood that a node i generate a *new unobserved edge* y_{ij} given it's degree d_i out of N_o observed edges can be estimated due to the conjugacy of priors within MMSB model. More formally let's first consider that the membership matrix Z associated with the observed degree is known. One can consider the given set $\mathcal{D} = \{d_i, Z\}$. Hence we can rewrite equation (**??**) as follows:

$$p(y_{ij}^{new} \mid \mathcal{D}, \alpha, \lambda) = \int_{fi} \int_{f_j} \int_\Phi p(f_i|\mathcal{D}, \alpha,)p(f_j|\mathcal{D}, \alpha)p(\Phi|\mathcal{D}, \lambda,) \tag{2.4}$$

$$\times \sum_{k<k'} p(y_{ij} \mid \phi_{kk'}) \; p(k \mid f_i)p(k' \mid f_j) df_i df_j d\Phi \tag{2.5}$$

One can notice the following useful identity here (we omit α here) by noticing that $di \perp\!\!\!\perp f_i \mid Z$, and that $f_i|\mathcal{D}$ and $\Phi|\mathcal{D}$ are respectively conjugate posteriors in the Dirichlet and beta distributions:

$$p(f_i|\mathcal{D}) = \frac{p(d_i, Z| \; fi)p(f_i)}{p(z, d_i)} = \frac{p(d_i|Z, f_i)p(Z|f_i)p(f_i)}{p(d_i|Z)p(Z)} = \frac{p(Z|f_i)p(f_i)}{p(Z)} \tag{2.6}$$

$$= Dir(f_i, \alpha + \mathbf{n_i}) \tag{2.7}$$

Where $\mathbf{n_i}$ is a k-rows vector representing the count of observed link or non-links such as node i take class k, therefore $n_{ik} = \sum_j \mathbf{1}(z_{i\rightarrow j} = k)$.

Similarly we obtains:

$$p(f_j|\mathcal{D}) = Dir(f_j, \alpha + \mathbf{n_j}) \tag{2.8}$$

Hence we have$n_{jk} = \sum_i \mathbf{1}(z_{i\leftarrow j} = k')$. Finally for the weight interaction we have that:

$$p(\phi_{kk'}|\mathcal{D}) = B(N_o - d_i^{k,k'} + \lambda_0, d_i^{k,k'} + \lambda_1) \tag{2.9}$$

With $d_i^{kk'} = \sum_j \mathbf{1}(y_{ij} = 1, z_{i\rightarrow j} = k, z_{i\leftarrow j} = k')$ and the joint distributions of weights interaction is:

$$p(\Phi|\mathcal{D}) = \prod_{k<k'} p(\phi_{kk'}|\mathcal{D}) \tag{2.10}$$

Thus, going back to equation (**??**), each factorial terms of the posteriors distribution have a term corresponding to the sum over all $\{k < k'\}_{k,k'=0}^{k,k'=K}$, that can be distributed on. Hence one can recognize that the predictive distribution of generating an edge between i and j is:

$$p(y_{ij}^{new} \mid \mathcal{D}, \alpha, \lambda) = \sum_{k<k'} \mathbb{E}[B((N_o - d_i)^{k,k'} + \lambda_0, d_i^{k,k'} + \lambda_1)]_{kk'} \mathbb{E}[Dir(\alpha + \mathbf{n_i})]_k \mathbb{E}[Dir(\alpha + \mathbf{n_j})]_{k'} \tag{2.11}$$

Finally we have that:

$$p(y_{ij}^{new} = 1 \mid \mathcal{D}, \alpha, \lambda) = \sum_{k<k'} \frac{d_i^{kk'} + \lambda_1}{N_o^{kk'} + \lambda_.} \frac{(n_{ik} + \alpha)(n_{jk'} + \alpha)}{(N + \alpha_.)^2} = \frac{1}{2} \hat{f}_i^T \hat{\Phi} \hat{f}_j \tag{2.12}$$

We see that the first term account for the node i popularity while the second term account for the similarity between i and j, and:

$$\hat{f}_{ik} = \frac{n_{ik} + \alpha}{N + \alpha_.} \tag{2.13}$$

$$\hat{f}_{jk} = \frac{n_{jk} + \alpha}{N + \alpha_.} \tag{2.14}$$

$$\hat{\phi}_{kk'} = \frac{d_i^{kk'} + \lambda_1}{N_o^{kk'} + \lambda_.} \tag{2.15}$$

One can notice that $d_i^{kk'} = d_i \frac{N_o^{kk',1}}{N_o^{kk'}} = d_i r_{kk'}$. This enable us to extract the degree. Morever let's define the following matrices:

- $S(i,j) = (\hat{f}_i \otimes \hat{f}_j)$ define the node similarity, depending on their class membership and $0 < S < 1$,
- $C = (\frac{1}{N_o^{kk'} + \lambda_.})_{kk'}$, a normalization factor for nodes interactions,
- $R = (r_{kk'})_{kk'}$ express the affinity between the two classes and $0 < R < 1$.

Finally the predictive likelihood can be expressed with those quantities as:

$$p(y_{ij}^{new} = 1 \mid \mathcal{D}, \alpha, \lambda) = \frac{1}{2}\ d_i\ sum(R \circ C \circ S(i,j)) + sum(\lambda_1 C \circ S(i,j)) \tag{2.16}$$

On the next sections we will study this distribution and analyse their meaning on the structure of generated networks.

2.3 Burstiness

2.4 Homophily

2.5 Small World

2.6 Sparsity