
1 Models and Properties

Nous définissons un modèle génératif selon 3 objets:

- Les hyperparamètres d'un modèle \mathcal{M}
- les paramètres aléatoires du modèle Θ . Dans les latent class models et particulièrement MMSB, on a $\Theta = \{F, \Phi, Z\}$
- Les observations Y qui sont les objets générés par le modèle.

Notre but est d'étudier le comportement de modèle génératif de réseaux. Pour cela on définit des propriétés adaptés à un cadre probabiliste. Nous affirmons qu'une propriété peut être considéré selon deux hypothèse distincte:

- propriété strong: propriété vrai et inhérente au modèle, on regarde ce que génère un réseaux étant donné ses paramètre aléatoires (intérêt pour l'ingénierie des paramètres...).
 $\Theta \longrightarrow Y$
- propriété weak: on ne connaît pas les paramètre aléatoires, seulement, les hyperparamètres. On raisonne selon ce schéma $\mathcal{M} \longrightarrow Y$

proposition1: Si une propriété est strong, alors elle est aussi "weak" vrai, car si elle est vrai pour un choix arbitraire de paramètres, ce choix tombe dans le tirage possible du modèles \mathcal{M} .

Dans la suite de cet écrit, nous proposons/définissons et vérifions le respect de propriétés fondamental des réseaux(preferential attach, homophily, small world, sparsity).

2 MMSB

...

3 Burstiness $\Delta_n p(y_i j = 1 | d_i > n) > 0$

i) Le réseau est trivialement **non strongly bursty**

ii) weakly ?

Mon raisonnement est le suivant:

1. on constate que la distribution du degré n'est pas exploitable et n'admet pas de forme close connue (see Neal)
2. on montre qu'en générant Z (ie Z est généré avec d_i , par le modèle), on a une expression [1].
3. on montre que cette expression permet de réécrire le burstiness
4. on montre que le burstiness est croissant si [1] est croissant
5. on montre que [1] est croissant avec n quelque soit les concentrations données par Z ...

De la définition du burstiness on obtient:

$$p(y_{ij} = 1 | d_i > n, \mathcal{M}) = \int_{\Theta} p(y_{ij} = 1 | \Theta) \frac{p(d_i > n | \Theta)}{p(d_i > n)} p(\Theta) d\Theta \quad (3.1)$$

Dans [*] on montre que distribution sur le degré n'a pas de forme connue (fait bien connue pour le $p(w)$ de LDA). Par contre, nous pouvons faire l'hypothèse que pour chaque valeur possible l du degré généré, soit accompagné de la matrice Z contenant les valeurs des classe d'attributions, pour chaque lien généré. Notons que nous ne faisons aucune hypothèse sur la distribution des classes d'attributions (c'est à dire de leur concentration, on dit simplement que chaque liens (ou non lien) ets associé à un couple kk'). la distribution est désormais implicitement conditionné par Z . Hors, étant donnée que le modèle est jointement échangeable [aldous, Hoff], on obtient:

$$p(y_{ij} = 1 | d_i > n, \mathcal{M}) = \int_{\Theta} p(y_{ij} = 1 | \Theta) \frac{\sum_{l=n}^N \binom{N}{l} p(d_i^N = l | \Theta)}{\int_{\Theta} \sum_{l=n}^N \binom{N}{l} p(d_i^N = l | \Theta)} p(\Theta) d\Theta \quad (3.2)$$

$$= \frac{\sum_l^N \binom{N}{l} \int_{\Theta} p(y_{ij} = 1 | \Theta) p(d_i^N = l | \Theta) p(\Theta) d\Theta}{\sum_l^N \binom{N}{l} \int_{\Theta} p(d_i^N = l | \Theta) p(\Theta) d\Theta} \quad (3.3)$$

$$= \frac{\sum_l^N a_l}{\sum_l^N b_l} \quad (3.4)$$

Interestingly, one can recognize that:

$$\frac{a_l}{b_l} = p(y_{ij} = 1 | d_i^N = l, Z) \quad (3.5)$$

We show in [*] that for an undirected network, the following holds:

$$pp(y_{ij} = 1 | d_i^N = l, Z) = \sum_{k < k'} \frac{d_i^{kk'} + \lambda_1}{N_o^{kk'} + \lambda_1} \frac{(n_{ik} + \alpha)(n_{jk'} + \alpha)}{(N + \alpha)^2} \quad (3.6)$$

$$= \frac{1}{2} \hat{f}_i^T \hat{\Phi} \hat{f}_j \quad (3.7)$$

We show that an other interesting parametrization, higlihting the degree dependance is:

$$p(y_{ij} = 1 | d_i^N = l, Z) = \frac{1}{2} l \text{ sum}(R \circ C \circ S(i, j)) + \text{sum}(\lambda_1 C \circ S(i, j)) \quad (3.8)$$

Where R , C and S are strictly positive matrix and their $\text{sum}(\cdot) < 1$. Moreover those matrix only depend on the class concentration whitin each node and more precisely see [*]:

- $S(i, j) = (\hat{f}_i \otimes \hat{f}_j)$ define the node similarity, depending on their class membership and $0 < S < 1$,
- $C = (\frac{1}{N_o^{kk'} + \lambda_1})_{kk'}$, a normalization factor for nodes interactions,

- $R = (r_{kk'})_{kk'}$ express the affinity between the two classes and $0 < R < 1$.

Thus one can write that $a_l = b_l g_l$, with g_l defined by (3.8).

Now let's define the following series:

$$a_n = \frac{\sum_{l=n}^N b_l g_l}{\sum_{l=n}^N b_l} \quad (3.9)$$

Given that b_l and g_l are positive, we show that if $\Delta_n g_n$ is increasing with n , then it implies that a_n is increasing with n :

proof:

$$\Delta_n a_n = \frac{\sum_{l=n+1}^N g_l b(l)}{\sum_{l=n+1}^N b(l)} - \frac{g_n b(n) + \sum_{l=n+1}^N g_l b(l)}{b(n) + \sum_{l=n+1}^N b(l)} \quad (3.10)$$

$$= \frac{b(n)}{\left(b(n) + \sum_{l=n+1}^N b(l)\right) \left(\sum_{l=n+1}^N b(l)\right)} \sum_{l=n+1}^N (g_l - g_n) b(l) \quad (3.11)$$

As we have $l \geq n$, it proves the hold.

Because g_l is linear¹ at a fixed class contraction state (corresponding to matrix Z) with the degree, it is straightforward that the networks satisfies the global preferential attachment.

Similarly, for the local preferential attachment we can assess that from the definition, when (i) and (j) are in class $c=k, k'$, we have simply that:

$$p(y_{ij} | d_{ic}^N = l, Z) = \frac{1}{2} l R_c C_c S(i, j)_c + \lambda_1 C_c S(i, j)_c \quad (3.12)$$

Again, we see that local preferential burstiness is true, furthermore as $R_c C_c S(i, j)_c + \lambda_1 C_c S(i, j)_c$ is a strictly increasing function of concentration of class assignment in c .

3.1 Discussion

* local preferential attachment inter-community!! (homophily vs stochastic equivalence). Pure power law function not possible (see expe)

* impact of global density

* Haldous hoover theorem and temporal networks sparsity and limit.

4 Discussions

* Relation with diaconis-Ylvisaker theorem, and how we generalize it, for mixed membership model in exponential family. (probabilistic.burstiness)

¹Note that this linearity is broken if the generation of new links would modify the concentration of links which is not the case in the assumptions of MMSB.