Final Project Proposal

Yi Chen

Teachers College Columbia University

Multidimensional Scaling, Clustering, and Network Models

Social network analysis refers to a collection of technique methods for describe the interactions among the individuals and to represent their relationship. In this final project, I will explore the latent space network model (Sweet and Adhikari, 2020) for social inference.

* **What is social inference**: social inference is also called spillover, contagion or diffusion): individuals become more similar to those with whom they interact or are most closely connected in their social network.
* **Example of social inference issue**: Adolescent’s delinquency behavior delinquency may be inferenced by the delinquency behavior delinquency of **friends** and also the delinquency behavior delinquency of the **popular student** (even they are *not* friend *directly*)
* **Benefit of using latent space model**: (1) identifying the underlying mechanism that generate the observed pattern in social network, (2) specifying the latent structure, which is able to identify the dependence/inference when there is no direct tie in the observed social network, and (3) the latent structure is also able to represent the unobserved latent co-determining variables. (4) in class, we discussed about how to make the social network plot based on the similarity/dissimilarity. This model instead, generate the latent space from the observed social network. And the latent space can be used to measure the distance/dissimilarity.
* **Model framework**: the social network structure by a adjacency / tie matrix,

, where is the value of the edge from actor to actor , which is usually binary value to represent the ties absent or present. The general latent space model that Sweet and Adhikari (2020)

proposed follows several assumptions:

1. *Conditional independence assumption*: the ties in the network are independent given the latent space variable.

2. *Monotonous assumption*: closer two nodes in the latent space, more likely the tie between the two nodes will be 1 (present).

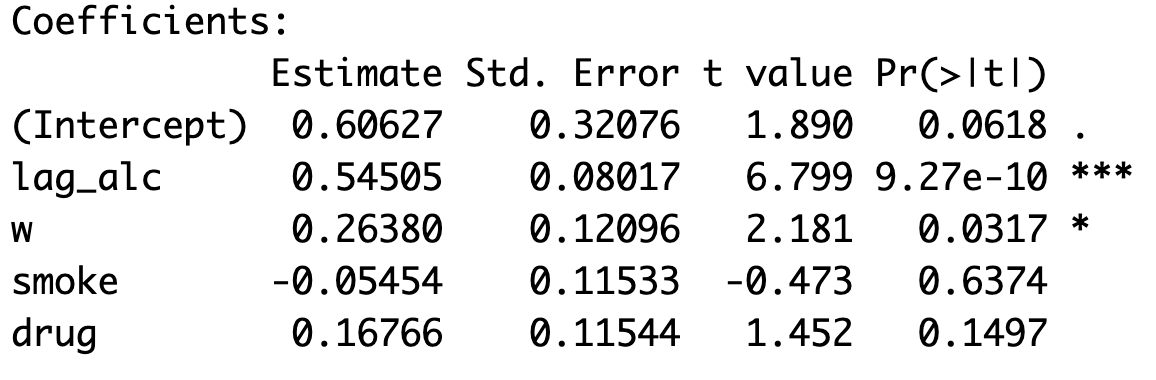
, where is the intercept term and represent the baseline probability of a tie for node in the network (overall density of the network), is a low -dimension vector of latent space location, is the distance between node and in the latent space, which can be calculate based on Euclidean distance. With these two basic assumptions, we can formulate the social inference model.

, is the collection of nodal outcomes measured at two different times and is a weight matrix of neighbor of the node (closer the neighbor is, bigger the weight it will have), is the intercept coefficient, is the effect of node outcome at the previous time on the outcome at the current time, and is the influence of the network on the outcome .

* **Example in R**
  + **Data Set**: Teenage Friends and Lifestyle Study data set (Michell 2000, Pearson and West 2003). <https://www.stats.ox.ac.uk/~snijders/siena/s50_data.htm>
  + Component:
    - Friendship network: Friendship network data and substance use were recorded for a cohort of *50* *female* pupils in a school in the West of Scotland.
    - The panel data were recorded over a three-year period starting in 1995, when the pupils were aged 13, and ending in 1997.
    - The friendship networks were formed by allowing the pupils to name up to twelve best friends.
    - Pupils were also asked about their attributes on *smoking (s), drug use (d), sport (sp),* and *alcohol use (a)*.
  + Model:

In this example, outcome of interest (Y) is attribute towards alcohol (a), other observed concurrent variables are their attribute towards drug (d) and smoke (s). Variables sport (sp) is used to capture the friendship network.

* + Result:



The social inference effect (w) is significant. While the biggest effect comes from the last years’ perception. Controlling the social inference and last years’ perception, the perception on smoke and drug are not significant any more.

* Note:
  + See ppt for more detailed information about the what I want to show during the presentation.
  + See R code for more detail about how to estimate the model.

References

Sweet, T. and Adhikari S. (2020). A Latent Space Network Model for Social Influence. *Psychometrika*, https://doi.org/10.1007/s11336-020-09700-x.

Xu, R. (2019) Estimating Social Influence Using Latent Space Adjusted Approach in R. *arXiv*

*preprint,* <https://arxiv.org/abs/1903.05999>.