

ALGORITMOS PARA LA DETECCIÓN DE COMUNIDADES

Nicolas Agote

Universidad de Buenos Aires, Argentina

agonico99@gmail.com

El modelo de bloques estocástico (Stochastic Block Model o SBM) [1] es un modelo de grafos aleatorios en el cual cada nodo cuenta con una etiqueta que determina sus conexiones. Más precisamente, se cuenta con una colección de nodos $V = \{1, \dots, n\}$ donde cada nodo pertenece a una de k comunidades $\{1, \dots, k\}$ y hay una arista entre nodos i y j de acuerdo a una distribución Bernoulli con parámetro que depende de las comunidades a las que i y j pertenecen, y cada arista se asigna de manera independiente. El objetivo de la detección de comunidades es dado un grafo aleatorio proviniendo de este modelo predecir cuáles nodos conforman las distintas comunidades con una probabilidad asintóticamente alta de tener una precisión adecuada.

En esta charla voy a presentar algunos algoritmos para realizar estas predicciones y relaciones entre ellos: algunos basados en paseos al azar en el grafo determinado por el modelo y grafos inducidos por él, y otros de la literatura como los que se pueden encontrar en [1], en [2] (basados en métodos espectrales) y en [3] (basados en encontrar una partición que realice una variante de corte mínimo para el grafo del modelo). En todo caso se estudia el modelo donde los parámetros de las distribuciones Bernoulli que determinan las aristas están en el orden $\frac{\log(n)}{n}$.

Este trabajo está siendo realizado para mi tesis de licenciatura, bajo la dirección de Inés Armendáriz (Universidad de Buenos Aires, Argentina).

Trabajo en conjunto con Inés Armendáriz (Universidad de Buenos Aires, Argentina), Pablo Ferrari (Universidad de Buenos Aires, Argentina), Florencia Leonardi (Universidade de São Paulo, Brasil) y Julio Rossi (Universidad de Buenos Aires, Argentina).

Referencias

- [1] Abbe, Emmanuel. (2017) "Community Detection and Stochastic Block Models". Disponible online en <https://arxiv.org/abs/1703.10146>
- [2] Mossel, Elchanan y Neeman, Joe y Sly, Allan. (2016) "Consistency thresholds for the planted bisection model" *Electronic Journal of Probability* 21, no. 21, 1–24. Disponible online en <https://arxiv.org/abs/1407.1591>
- [3] Hein, Matthias y Bühler, Thomas. (2010) "An Inverse Power Method for Nonlinear Eigenproblems with Applications in 1-Spectral Clustering and Sparse PCA". Disponible online en <https://arxiv.org/abs/1012.0774>