A Matrix Factorization Framework for Community Detection under the Degree-Corrected Block Model

Alexandra Dache*[1], Nicolas Gillis[1], and Arnaud Vandaele[1]

[1]Université de Mons – Belgium

Abstract

Community detection is an essential tool in data analysis. A widely used approach is clustering based on a graph model (block modeling), which allows for the identification of various underlying structures. Block models partition nodes into clusters that exhibit similar connection patterns, facilitating structural analysis and providing a deeper understanding of the network. The Degree-Corrected Block Model (DCBM) is a well-established model that generalizes the Stochastic Block Model by addressing node degree heterogeneity. DCBM owes its success to its simplicity and the variety of network structures it can model. While numerous inference methods exist to estimate the parameters of the DCBM from the graph, they are all heuristic-based. As a result, the quality of the results heavily depends on initialization, which is often performed randomly. We propose a new approach: instead of relying on a probabilistic modeling framework, we reformulate the DCBM parameter estimation problem as a nonnegative matrix factorization (NMF) problem. This development led us to devise an effective method based on separable NMF for parameter initialization, a critical step in the inference process. We show that the NMF model derived from the DCBM and dedicated algorithm yield excellent results on benchmark networks, both real and synthetic, and that our initialization strategy significantly improves the convergence of existing inference methods for the DCBM.

*Speaker