



Working Group on Statistical Learning Seminar

Title: A Dirichlet stochastic block model for composition-weighted network

Speaker: Iuliia (Yulia) Promskaia (UCD)

Date: Thu 15th February 2024 at 3:00PM

Location: E0.32 (beside Pi restaurant)

Abstract: In this work we propose a Dirichlet stochastic block model (DirSBM) for clustering nodes in networks with compositional edge weights. The DirSBM models the compositional weight vectors directly using a Dirichlet mixture, with parameters determined by the cluster labels of the sender and receiver nodes. We address the inference problem via a recently proposed hybrid maximum likelihood approach by Marino Pandolfi (2022) for stochastic block models. This inferential procedure allows us to derive a variant of the classification expectation-maximisation algorithm by utilising a working independence assumption, expressing the conditional distribution of compositional weights for each node as a function of the fixed cluster labels of the remaining nodes in the network. An integrated completed likelihood criterion is derived for the purpose of model selection. An alternative approach to clustering nodes in composition-weighted networks based on a mapping to the Euclidean space is provided for comparison. The proposed DirSBM is tested in a number of synthetic data experiments, assessing the effect of various initialisation strategies, and evaluating the model's performance in relation to cluster allocation recovery, model selection, and accuracy of parameter estimation. The model is also illustrated in application to two real-world data sets: a network of student exchanges between European countries within the Erasmus programme, and a bike sharing network for the city of London.