Matteo Ruffini
Universitat Politècnica de Catalunya

Clustering patients with Tensor Decomposition

Clustering patients in groups with similar clinical profiles is a strategic activity for a modern healthcare systems. This, typically requires to deal with high-dimensional categorical data, a framework where most of the traditional clustering algorithms perform poorly. A viable alternative is to use a clustering procedure based on generative latent variable models, a task that requires fitting the parameters of the models in question. In recent times, learning the parameters of latent variable models has become a popular research branch, due to the nice tensor structure of their low-order observable moments. Methods based on tensor decomposition have emerged as effective and efficient alternatives to the existing heuristics and their applications are rapidly growing.
In this talk, we will see how tensor decomposition can be used to fit the parameters of latent variable models, focusing on some special cases of unlabeled mixture models and on how they can be used as a clustering tool. We will then see the application of this clustering procedure to medical data, using it to automatically generate clusters of patients with homogeneous clinical profiles.