

An Updated EM Algorithm for Classification in Protein Interaction Networks

Hanen Ben Hassen¹, Aida Jarraya²

¹ Physique-Mathématiques et Applications. Faculty of sciences of Sfax,
Tunisia

hanenbenhassen@yahoo.fr

² Laboratory of Probability and Statistics, Faculty of Sciences of Sfax,
Sfax University. B.P. 1171 Sfax. Tunisia

jarrayaaida@yahoo.fr

Keywords: Biomolecular networks, Classification, EM Algorithm, Graphical Model, Mixture of Gaussian Distributions.

We provide an updated Expectation-Maximization (UEM) algorithm to estimate parameters of finite Gaussian mixture distributions in Bayesian networks. The UEM algorithm is composed by three steps: an expectation (E) and a maximization (M) steps, akin to the EM algorithm, where a created function for log-likelihood is evaluated using the current estimate for the parameter after that the parameters estimated are maximized. Given a graph specifying the relationship between nodes and based on Lauritzen formula, the third step (U) updates the M-estimates. We demonstrate that UEM algorithm is useful to system biology data, by considering protein interaction networks where arcs represent probabilistic relationships (interaction) between nodes (proteins). The UEM algorithm provides the classification of the data to each Gaussian distribution cluster. We apply our algorithm on a simulation study of the EGFR protein interaction network.