Nonparametric Estimation in Multivariate Finite Mixture Models

Tatiana Benaglia

Pennsylvania State University, USA

Abstract

We propose an algorithm for nonparametric estimation for finite mixtures of multivariate random vectors that is not, but that strongly resembles, a true EM algorithm. The vectors are assumed to have conditionally independent coordinates, but otherwise their density functions are completely unspecified. Our algorithm is much more flexible and easily applicable than existing algorithms in the literature; it can be extended to any number of mixture components and any number of coordinates of the multivariate observations. Thus it may be applied even in situations where the model is not identifiable, so care is called when it is difficult to establish identifiability conclusively. Our algorithm yields much smaller mean integrated squared errors than an alternative algorithm in a simulation study. In another example using a real dataset, it provides new insights that extend previous analyses.

Keywords: EM algorithm, kernel density estimation, multivariate mixture, nonparametric mixture.