

1153-62-105

Raaz Dwivedi* (raaz.rsk@berkeley.edu). *Singularity, Misspecification and Convergence of Expectation Maximization.*

A line of recent work has characterized the behavior of the EM algorithm in favorable well-specified settings in which the population likelihood is locally strongly concave around its maximizing argument. Examples include suitably separated Gaussian mixture models and mixtures of linear regressions. We consider instead over-specified settings in which the number of fitted components is more than that in the true distribution. In such cases, the Fisher information matrix can be singular, and consequently, the MLE based on n samples can have slower than $n^{-\frac{1}{2}}$ rate of convergence. Focusing on the simple setting of a two-component mixture fit to a Gaussian distribution, we study the behavior of the EM algorithm to estimate the location and/or scale parameter. When scale parameter is known, we show that EM's fixed point has a much lower $n^{-\frac{1}{4}}$ accuracy. When the scale parameter is also estimated, EM slows down further and the estimation error has a scaling of order $n^{-\frac{1}{8}}$. Analysis of the singular cases requires the introduction of some novel techniques: in particular, we make use of a careful form of localization in the associated empirical process and develop a recursive argument to progressively sharpen the statistical rate. (Received August 19, 2019)