Estimating densities with nonlinear support using Fisher-Gaussian kernels

Minerva Mukhopadhyay

Indian Institute of Technology Kanpur
E-mail: minervamukherjee@gmail.com

Abstract: Current tools for multivariate density estimation struggle when the density is concentrated near a nonlinear subspace or manifold. Most approaches require choice of a kernel, with the multivariate Gaussian kernel being the most commonly used one. Although heavy-tailed and skewed extensions have been proposed, such kernels cannot capture curvature in the support of the data. Hence, a very large number of kernels may be needed to provide an adequate fit to many datasets. This leads to poor performance unless the sample size is very large relative to the dimension of the data, even in toy problems. With this motivation, we propose a novel generalization of the Gaussian distribution, which includes an additional curvature parameter. We refer to the proposed class as Fisher-Gaussian (FG) kernels, since they arise by sampling data from a von Mises-Fisher density on the sphere and adding Gaussian noise. The FG density has an analytic form, and is amenable to straightforward implementation within Bayesian mixture models using Markov chain Monte Carlo. We provide asymptotic theory on posterior concentration, and illustrate gains relative to current methods on simulated and real data applications.