High-dimensional log-concave density estimation

Richard Samworth

University of Cambridge
E-mail: r.samworth@statslab.cam.ac.uk

Abstract: We tackle the problem of high-dimensional nonparametric density estimation by taking the class of log-concave densities on $\mathbb{R}^p$ and incorporating within it symmetry assumptions, which facilitate scalable estimation algorithms and can mitigate the curse of dimensionality. Our main symmetry assumption is that the super-level sets of the density are K-homothetic (i.e. scalar multiples of a convex body $K \subseteq \mathbb{R}^p$). When K is known, we prove that the K-homothetic log-concave maximum likelihood estimator based on n independent observations from such a density has a worst-case risk bound with respect to, e.g., squared Hellinger loss, of $O(n^{-4/5})$, independent of p. Moreover, we show that the estimator is adaptive in the sense that if the data generating density admits a special form, then a nearly parametric rate may be attained. We also provide worst-case and adaptive risk bounds in cases where K is only known up to a positive definite transformation, and where it is completely unknown and must be estimated nonparametrically.