Kernel Density Construction by Variational Method

Kōsaku Takanashi
Riken AIP

April, 2019, Keio University

Non-parametric kernel density estimation methods are now popular and in wide use with great success in statistical applications. Kernel density estimates are commonly used to display the shape of a data set without relying on a parametric model, not to mention the exposition of skewness, multimodality, dispersion, and more. The monographs of [2] and [3] are good references.

However, kernel estimators suffer from **boundary bias** when the support of the function to be estimated is compact space due to the fact that most kernels do not take into account specific knowledge about the domain of the data. Boundary bias seriously affect the overall performance of the estimator. The Gaussian kernel density estimator lacks **local adaptivity**, resulting in a large sensitivity to outliers, the presence of spurious bumps, and in a tendency to flatten the peaks and valleys of the density.

[1] has introduced an adaptive kernel density estimation method based on the smoothing properties of linear diffusion processes. The resulting diffusion estimator unifies many of the existing ideas about adaptive smoothing, and is consistent at boundaries. [1] interpret the kernel from which the estimator is constructed as the transition density of a diffusion process. They construct a general linear diffusion process that has a given limiting and stationary probability density which is selected to be either a pilot density estimate or

a prior density that the statistician believes represents the information about the data prior to observing the available empirical data. This approach leads to substantially reduced asymptotic bias and mean square error and deals well with boundary bias. Moreover, [1] introduces an improved plug-in bandwidth selection method that completely avoids the so-called normal reference rules that have adversely affected the performance of plug-in methods.

In this thesis, we extend their proposed method in two ways. First, we extend their proposed diffusion kernel method to kernel density estimators based on Lévy processes, which have the diffusion estimator as a special case. The kernels constructed via a Lévy process could be tailored for data for which smoothing with the diffusion estimator is not optimal. Such cases arise when the data is a sample from a heavy-tailed distribution. It is true for compactly supported and not heavy-tailed distributions, the diffusion kernel methods are a good estimate of the corresponding probability density function. But for the heavy-tailed distribution, the diffusion methods provide misleading peaks in the 'tail' domain or oversmoothes the 'body' of the density.

Second, we consider an asymptotics of the estimated diffusion differential operator that has a random fluctuate due to the estimated pilot density. Then, constructing PDE kernel by an adaptive method proceed to a variational problem; when the estimated diffusion differential operator converges to some pseudo limit operator in some topological sense, dose a kernel induced the diffusion equation with the estimated diffusion differential operator also converge to some limit kernel induced by the pseudo limit operator? This problem induces a variational problem, and in fact can be addressed by a straightforward application of Mosco convergence of Dirichlet form.

References

[1] Zdravko I Botev, Joseph F Grotowski, Dirk P Kroese, et al. Kernel density estimation via diffusion. *The annals of Statistics*, 38(5):2916–2957, 2010.

[2] BW Silverman. Density estimation for statistics and data analysis. 1986.

[3] MP Wand and MC Jones. Kernel smoothing. 1995. *Chapman&Hall, London*, 1995.