Learning Adaptive Multiscale Approximations to Data and Functions near Low-Dimensional Sets

Mauro Maggioni
Department of Mathematics, Electrical and Computer Engineering, and Computer Science, Duke University, Durham, N.C., U.S.A.
Email: mauro.maggioni@duke.edu

In the setting where a data set in $\mathbb{R}^D$ consists of samples from a probability measure ρ concentrated on or near an unknown d-dimensional set $\mathcal{M}$, with D large but $d \ll D$, we consider two sets of problems: geometric approximation of $\mathcal{M}$ and regression of a function f on $\mathcal{M}$. In the first case we construct multiscale low-dimensional empirical approximations of $\mathcal{M}$, which are adaptive even when $\mathcal{M}$ has geometric regularity that may vary at different locations and scales, and give performance guarantees. In the second case we exploit these empirical approximations to construct multiscale approximations of f on $\mathcal{M}$, which adapt to the unknown regularity of f even when this varies at different scales and locations, and prove guarantees that show that we attain the same learning rates as if f was defined on a Euclidean domain of dimension d, instead of an unknown manifold $\mathcal{M}$. All algorithms have complexity $O(n \log n)$.