

1145-62-1044

Didong Li* (didongli@math.duke.edu), 120 Science Drive, 117 Physics Building, Campus Box 90320, Durham, NC 27708, and **Minerva Mukhopadhyay** and **David B Dunson**. *Efficient Manifold Approximation with Spherelets*.

Data lying in a high dimensional ambient space are commonly thought to have a much lower intrinsic dimension. In particular, the data may be concentrated near a lower-dimensional subspace or manifold. There is an immense literature focused on approximating the unknown subspace, and in exploiting such approximations in clustering, data compression, and building of predictive models. Most of the literature relies on approximating subspaces using a locally linear, and potentially multiscale, dictionary. In this talk, we propose a simple and general alternative, which instead uses pieces of spheres, or spherelets, to locally approximate the unknown subspace. Theory is developed showing that spherelets can produce lower covering numbers and MSEs for many manifolds. We develop spherical principal components analysis (SPCA). Results relative to state-of-the-art competitors show gains in ability to accurately approximate the subspace with fewer components. In addition, unlike most competitors, our approach can be used for data denoising and can efficiently embed new data without retraining. The methods are illustrated with standard toy manifold learning examples, and applications to multiple real data sets. (Received September 18, 2018)