DEPARTMENT OF STATISTICS

Xiaohui Chen
Department of Statistics
University of Illinois at Urbana-Champaign

A Diffusion Perspective of Manifold Clustering

Wednesday, September 2nd, 2020
11:45 – 12:45PM EST

Zoom Meeting: 983 1453 7803
Password: 491555
Virtual Coffee session before the seminar at 11:30AM EST

Abstract: We introduce the diffusion K-means clustering method on Riemannian submanifolds, which maximizes the within-cluster connectedness based on the diffusion distance. The diffusion K-means constructs a random walk on the similarity graph with vertices as data points randomly sampled on the manifolds and edges as similarities given by a kernel that captures the local geometry of manifolds. The diffusion K-means is a multi-scale clustering tool that is suitable for data with non-linear and non-Euclidean geometric features in mixed dimensions. Given the number of clusters, we propose a polynomial-time convex relaxation algorithm via the semidefinite programming (SDP) to solve the diffusion K-means. In addition, we also propose a nuclear norm regularized SDP that is adaptive to the number of clusters. In both cases, we show that exact recovery of the SDPs for diffusion K-means can be achieved under suitable between-cluster separability and within-cluster connectedness of the submanifolds, which together quantify the hardness of the manifold clustering problem. We further propose the localized diffusion K-means by using the local adaptive bandwidth estimated from the nearest neighbors. We show that exact recovery of the localized diffusion K-means is fully adaptive to the local probability density and geometric structures of the underlying submanifolds. Joint work with Yun Yang (UIUC).

Bio: Xiaohui Chen received a Ph. D. in Electrical and Computer Engineering in 2013 from the University of British Columba (UBC), Vancouver, Canada. He was a post-doctoral fellow at the Toyota Technological Institute at Chicago (TTIC), a philanthropically endowed academic computer science institute located on the University of Chicago campus. In 2013 he joined the University of Illinois at Urbana-Champaign (UIUC) as an Assistant Professor of Statistics. He is an Associate Professor of Statistics at UIUC since 2019 and a member of Discovery Partners Institute (DPI) since 2020. He held Visiting Faculty position in the Institute for Data, Systems, and Society (IDSS) at Massachusetts Institute of Technology (MIT) in 2019-2020. He received numerous notable awards, including an NSF CAREER Award in 2018, an Arnold O. Beckman Award at UIUC in 2018, an Outstanding Young Researcher Award from the International Chinese Statistical Association (ICSA) in 2019, an Associate appointment in the Center for Advanced Study at UIUC in 2020-2021, and a Simons Fellowship in Mathematics from the Simons Foundation in 2020-2021. His teaching was recognized three times by the University of Illinois List of Teachers Ranked as Excellent by Their Students.