

学术报告

Stochastic Approximation for Feature Extraction:
Nonconvex Optimization, Online Learning, and
Diffusion Approximation

李骏驰 博士
美国普林斯顿大学

报告时间：2018 年 4 月 2 日（周一）下午 4:00-5:00

报告地点：天津大学 6 号楼 111 教室

报告摘要： Statistical dimension-reduction and feature-extraction procedures such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), Partial Least Squares (PLS) and Canonical Correlation Analysis (CCA) present significant computational challenges with large-scale data. Online algorithms that updates the estimator by processing streaming data on-the-fly are of tremendous practical and theoretical interests. In this talk, we formulate these statistical procedures as nonconvex statistical optimization problem with nonconvex statistical structures, and view these online algorithm as corresponding stochastic approximation methods. Using standard tools of martingale concentration inequalities, we for the first time obtain the finite-sample error bound of $O(\sqrt{d/N})$. We prove that (i) for PCA, such bound is known to closely match the minimax information lower bound for PCA (L. et al., 2015 Mathematical Programming; L. et al., 2017 NIPS); (ii) for (tensor formulation of) ICA, such bound is consistent with the computational lower bound for spiked tensor model (L., Wang & Liu, 2016 NIPS; L. et al., 2017; Wang, Yang, L. & Liu, 2016); (iii) for PLS and CCA, novel online location-dependent stochastic gradient method is proposed to achieve this bound. Time permitting, we further brief the recent progresses on the diffusion approximation methods for generic first-order algorithms. We show that in the continuous-time framework, first-order stochastic gradient method (Hu & L., 2017+) as well as stochastic heavy-ball method (Hu, L. & Su, 2017+) fastly avoid all saddle points within a time complexity that is inverse proportional to the stepsize.

欢迎大家参加！