



“MATH for Machine Learning” Seminar

Sharper Generalization Guarantees for Asynchronous SGD: Beyond Lipschitzness, Smoothness and Data Homogeneity

Mr. Yufeng Xie

HKU

Abstract

Asynchronous stochastic gradient descent (ASGD) is widely adopted in distributed and federated learning. We develop a sharp generalization analysis for ASGD by leveraging the concept of on-average model stability with respect to sub-dataset. For convex and smooth objectives, we establish stability and excess risk bounds under minimal assumptions, removing assumptions on Lipschitz continuity, bounded noise, bounded parameter or data domains, while allowing randomly partitioned data and arbitrary delays. Our bounds are optimistic and explicitly characterize the impact of worker participation, recovering the minimax-optimal rate $O(1/\sqrt{mn})$ in balanced regimes where mn denotes the sample size and implying fast rates under low-noise conditions. We further extend the analysis to non-smooth objectives with Hölder-continuous gradients and to heterogeneous data settings via random ASGD, obtaining non-vacuous excess risk guarantees in both settings. Experimental results support our theoretical findings.

Date:	September 24, 2026 (Thursday)
Time:	3:00 pm – 4:00 pm
Venue:	Room 210, Run Run Shaw Building HKU