FUDAN UNIVERSITY

Applied Math Ph.D. Seminar

A General Framework for Decentralized Stochastic Subgradient-based Methods in Nonsmooth Nonconvex Optimization

Speaker: Siyuan Zhang (Chinese Academy of Sciences)
Time: 2024-11-28, 16:10 to 17:00
Location: Rm 1801, Guanghua East Tower
Advisor: Xin Liu (Chinese Academy of Sciences)

Abstract: Benefiting from the ability to reduce communication overhead and maintain the privacy of local agents, decentralized optimization methods have come to the forefront in distributed learning, especially in the decentralized training of DNN. However, the widely employed ReLU activation function leads to the nonsmoothness without Clarke regularity of the loss function in the training of neural networks, which leads to a gap between existing theoretical analysis and implementation in real-world training tasks. In this poster, we concentrate on decentralized optimization problems with nonsmooth nonconvex objective functions. We introduce a unified framework, named DSM, to analyze the global convergence of decentralized stochastic subgradient-based methods, by establishing that the generated sequence asymptotically approximates the trajectories of its associated differential inclusion. We show our proposed framework covers a wide range of existing efficient decentralized subgradient-based methods, such as DSGD, DSGD with gradient-tracking technique (DSGD-T), and DSGD with momentum (DSGD-M). In addition, we introduce the sign map to regularize the update directions in DSGD-M, and show it is enclosed in our proposed framework. Preliminary numerical experiments verify the validity of our developed theoretical results and exhibit flexibility and potential in developing new decentralized subgradient-based algorithms.