

STRUCTURAL INVARIANT MANIFOLDS OF NEURAL NETWORK TRAINING DYNAMICS

Speaker: Jiajie Zhao 赵佳杰

SJTU (上海交通大学)

Time： Wednesday, Dec. 2nd, 16:00
Venue: Room 102, SCMS

Abstract:

In this talk, I will discuss how neural network architecture constrains gradient flow dynamics. We introduce structural invariant manifolds (SIMs), which are invariant under gradient flow independently of the training data and loss function. Using geometric control theory, we show that SIMs are precisely unions of orbits generated by a family of vector fields determined by the network architecture. We also identify a hierarchy of symmetry-induced SIMs in deep neural networks and prove that all SIMs of generic two-layer networks are symmetry-induced. Finally, we study diagonal linear networks under infinitesimal initialization and show how SIMs shape saddle-to-saddle dynamics and explain the resulting implicit bias toward modified L1-norm minimization.