



Optimization and Machine Learning Seminar

On Generalization and Implicit Rank Restriction in Deep Matrix Completion and Neural Networks

Prof. Antoine Ledent

Singapore Management University, Singapore

Abstract

Low-rank structure is ubiquitous in machine learning methods. It can be either imposed implicitly or arise as a byproduct of the optimization constraints or algorithm. For instance, several early matrix completion methods constrain the rank of the solution to a fixed value, and some deep learning methods such as Low Rank Adaptation (LoRA) consider low-rank updates to transformer weights. In this talk, we will begin with an introduction to the matrix completion problem and the related literature. Then, we will evaluate the effects of the Schatten quasi norm restrictions on generalization in two settings: matrix completion and neural networks. In particular, we will present generalization bounds which capture the approximate low-rank structure associated with low Schatten quasi norm matrices. In both matrix completion and deep neural networks, the bounds converge to the results one might expect from a parameter counting argument as the depth increases, thereby capturing the rank-sparsity inducing effect of depth in both cases. We also discuss extensions to neural matrix factorization models and directions for future work.

Date: March 6, 2025 (Thursday)

Time: 4:00 pm - 5:00 pm

Venue: Room 210, Run Run Shaw
Building HKU