



SLING Efficient algorithms for sustainable machine learning

Mon, May 30th, 2022, 4:00 p.m., DIBRIS, via Dodecaneso 35, Genova.

Statistical Learning & Optimization

The implicit bias of gradient descent and its applications

Abstract.

In deep learning it is common to overparameterize neural networks, that is, to use more parameters than training samples. Quite surprisingly, training these networks via (stochastic) gradient descent leads to models that generalize very well, while classical statistics would suggest overfitting. In this talk, we theoretically analyze the behaviour of vanilla gradient flow/descent in two simplified settings: (i) matrix factorization (which can be seen as training linear neural networks without bias) and (ii) sparse recovery. Whereas in (i) the iterates follow a path of low (effective) rank, in (ii) the limit (approximately) minimizes the ℓ_1 -norm among all possible solutions (under very mild assumptions on the measurement matrix). We conclude the talk with some recent insights into how the implicit bias can be used for solving classical problems like non-negative least squares (NNLS).

This talk is based on joint work with Hung-Hsu Chou, Carsten Gieshoff, Holger Rauhut, and Claudio Verdun.

Speaker

Johannes Maly

Catholic University of Eichstaett-Ingolstadt



I am a lecturer at the Catholic University of Eichstaett/Ingolstadt in Germany. Previously, I have been working as a postdoctoral researcher at RWTH Aachen. I obtained my PhD in 2019 from TUM in Munich. My research focuses on recovery of multi-structured signals, covariance estimation, approximation properties of neural networks, and the implicit bias of gradient descent (and particularly in understanding the influence of coarse quantization in all of these).

Host: Silvia Villa

contact: malga.unige@gmail.com