Analysing Implicit Bias/Regularisation via Invariant, Bregman Divergence, and Normalisation

Hung-Hsu Chou

Ludwig-Maximilians-Universität München, Germany

chou@math.lmu.de

Out of infinitely many possible explanations to the observed data, machines are often capable of approaching the ones with "good" properties, such as generalisability or compressibility, even when those properties are not explicitly specified in the training algorithms. Such phenomenon, known as the implicit bias/regularisation, has been intensively studied in the last decade and stimulated significant development of fundamental understanding in machine learning.

My presentation focuses on analysing the implicit bias from different tools and perspectives: invariants, Bregman divergence, and normalisation. I use invariants, quantities that do not change in time, to identify the region where the training process lies. I further analyse the process via Bregman divergence, a generalisation of distance, that is embedded in the algorithm and use it to characterise the end result of the training. Moreover, I discover that additional normalisation in the algorithm strengthen its robustness with respect to the change in initialization, and consequently increase the numerical efficiency.

I am currently focusing on extending tools and concepts to a wider range of models, such as linear networks and multilayer perceptrons. I am also interested in neural tangent kernel, neural collapse, spectral bias, and related fields.

Joint work with Johannes Maly (Ludwig-Maximilians-Universität München, Germany), Holger Rauhut (Rheinisch-Westfälische Technische Hochschule Aachen, Germany), Dominik Stöger (Katholische Universität Eichstätt-Ingolstadt, Germany), Cristian Vega (University of Genoa, Italy) and Rachel Ward (University of Texas at Austin, United States of America).