Title: Optimization bias in linear convolutional networks

Abstract: In the modern practice of machine learning, especially in deep learning, many successful models are highly overparameterized with far more trainable parameters compared to the number of training examples. Consequently the resulting optimization objective has many solutions that perfectly fit the training data, most of which will perform poorly on new samples. It is now evident that the implicit inductive bias from optimization algorithms like (stochastic) gradient descent play a crucial role in learning meaningful models from such ill-posed objectives.

In this talk I will briefly overview some exciting results on understanding the inductive bias from optimization algorithms. In particular, I will talk about a recent work that shows how in learning even simple linear models using gradient descent on different neural network architectures, such as fully connected or convolutional networks, lead to very different and interesting optimization biases.

Based on following joint work with Jason Lee, Daniel Soudry, and Nati Srebro;
Implicit Bias of Gradient Descent on Linear Convolutional Networks (NeurIPS 2018)