



THE UNIVERSITY OF
CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

TENGYU MA

Department of Computer Science
Stanford University
Facebook AI Research

Algorithmic Regularization in Over-parameterized Matrix Recovery
and Neural Networks with Quadratic Activations

MONDAY, May 21, 2018 at 4:30 PM
Eckhart 133, 5734 S. University Avenue
Refreshments before the seminar at 4:00PM in Jones 111

ABSTRACT

Over-parameterized models are widely and successfully used in deep learning, but their workings are far from understood. In many practical scenarios, the learned model generalizes to the test data, even though the hypothesis class contains a model that completely overfits the training data and no regularization is applied. In this talk, we will show that such phenomenon occurs in over-parameterized matrix recovery models as well, and prove that the gradient descent algorithm provides additional regularization power that prevents the overfitting. The result can be extended to learning one-hidden-layer neural networks with quadratic activations. The key insight here is that gradient descent prefers searching through the set of low complexity (that is, low-rank) models first, and converges to a low complexity model with a good training error if such a model exists.

Based on joint work with Yuezhi Li and Hongyang Zhang.
<https://arxiv.org/abs/1712.09203> (This is joint research with Veronika Rockova, Paul Rosenbaum, Ville Satopaa and Jeffrey Silber).

For further information and inquiries about building access for persons with disabilities, please contact Jonathan Rodriguez at 773.702.8333 or send him an email at jgrodriquez@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following website:
<https://lists.uchicago.edu/web/subscribe/statseminars>.