



SPRING 2019

RENSSELAER POLYTECHNIC INSTITUTE

DEPARTMENT OF MATHEMATICAL SCIENCES COLLOQUIUM/RTG SEMINAR

"Neural networks as interacting particle systems: understanding global convergence of parameter optimization dynamics"

Abstract: The performance of neural networks on high-dimensional data distributions suggests that it may be possible to parameterize a representation of a target high-dimensional function with controllably small errors, potentially outperforming standard interpolation methods. We demonstrate, both theoretically and numerically, that this is indeed the case. We map the parameters of a neural network to a system of particles relaxing with an interaction potential determined by the loss function. This mapping gives rise to a deterministic partial differential equation that governs the parameter evolution under gradient descent dynamics. We also show that in the limit that the number of parameters n is large, the landscape of the mean-squared error becomes convex and the representation error in the function scales $\text{link } n^{-1}$. In this limit, we prove a dynamical variant of the universal approximation theorem showing that the optimal representation can be attained by stochastic gradient descent, the algorithm ubiquitously used for parameter optimization in machine learning. This conceptual framework can be leveraged to develop algorithms that accelerate optimization using non-local transport. I will conclude by showing that using neuron birth/death processes in parameter optimization guarantees global convergence and provides a substantial acceleration in practice.

Grant Rotskoff, Courant Institute, NYU

Monday, April 8, 2019 4-5pm

Amos Eaton 214

Host: Peter Kramer

Refreshments served 3:30-4pm Amos Eaton 4th Floor Lounge