

CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday – January 3, 2020

**Paper review: ANTISYMMETRICRNN: A DYNAMICAL SYSTEM VIEW
ON RECURRENT NEURAL NETWORKS**

Zhen Zhang

Recurrent neural networks have gained widespread use in modeling sequential data. Learning long-term dependencies using these models remains difficult though, due to exploding or vanishing gradients. In this paper, we draw connections between recurrent networks and ordinary differential equations. A special form of recurrent networks called the AntisymmetricRNN is proposed under this theoretical framework, which is able to capture long-term dependencies thanks to the stability property of its underlying differential equation. Existing approaches to improving RNN trainability often incur significant computation overhead. In comparison, AntisymmetricRNN achieves the same goal by design. We showcase the advantage of this new architecture through extensive simulations and experiments. AntisymmetricRNN exhibits much more predictable dynamics. It outperforms regular LSTM models on tasks requiring long-term memory and matches the performance on tasks where short-term dependencies dominate despite being much simpler.