

NUMERICAL ANALYSIS AND SCIENTIFIC COMPUTING
SEMINAR

*Numerical Linear Algebra Methods in Recurrent Neural
Networks*

Qiang Ye
University of Kentucky

Abstract: Deep neural networks have emerged as one of the most powerful machine learning methods. Recurrent neural networks (RNNs) are special architectures designed to efficiently model sequential data by exploiting temporal connections within a sequence and handling variable sequence lengths in a dataset. However, they suffer from so-called vanishing or exploding gradient problems. Recent works address this issue by using a unitary/orthogonal recurrent matrix. In this talk, we will present some numerical linear algebra based methods to improve RNNs. We first introduce a simpler and novel RNN that maintains orthogonal recurrent matrix using a scaled Cayley transform. We then develop a complex version with a unitary recurrent matrix that allows direct training of the scaling matrix in the Cayley transform. We further extend our architecture to use a block recurrent matrix with a spectral radius bounded by one to effectively model both long-term and short-term memory in RNNs. Our methods achieve superior results with fewer trainable parameters than other variants of RNNs in a variety of experiments.

Friday, October 9, 2020, 2:40 pm
<https://emory.zoom.us/j/95900585494>

MATHEMATICS
EMORY UNIVERSITY