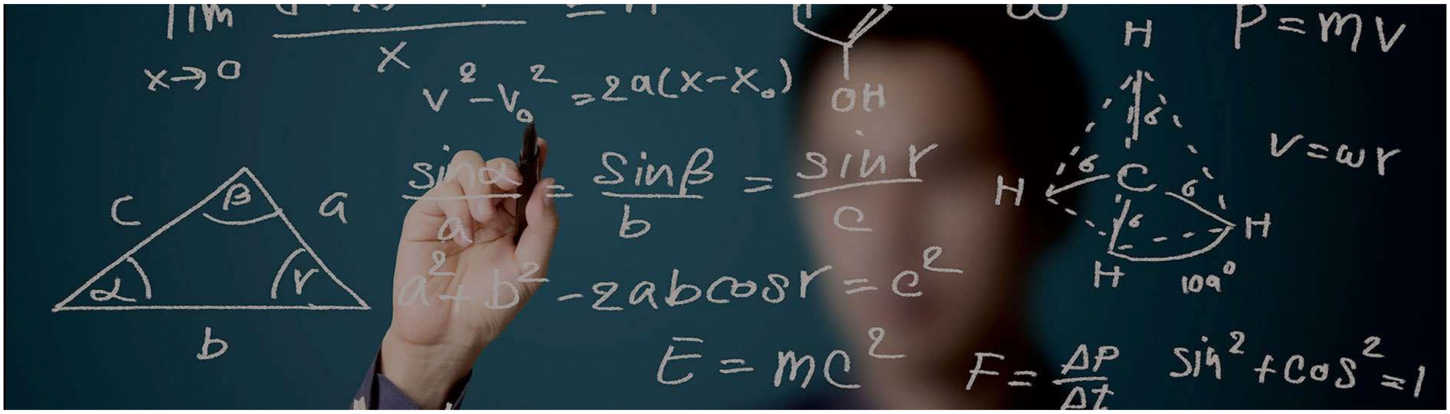

Spring 2024

RENSSELAER POLYTECHNIC INSTITUTE

DEPARTMENT OF MATHEMATICAL SCIENCES COLLOQUIUM

Davoud AtaeeTarzanagh (University of Pennsylvania)
January 25, 2024 - 4pm
PITTS 4206

Understanding the Optimization Geometry of the Attention Mechanism

Attention mechanism is a central component of the transformer architecture, which has led to the phenomenal success of large language models. Its effectiveness in capturing long-range dependencies is pivotal for tasks in natural language processing, computer vision, and reinforcement learning. Despite its practical success, the theoretical foundations of the attention mechanism, particularly its nonconvex optimization dynamics, remain largely unexplored.

In this talk, we provide a mathematically rigorous understanding of how the attention mechanism processes input sequences, such as sentences composed of words, despite the high-order interactions among the elements in the sequence. Our approach involves analyzing the convergence properties of gradient descent during the training of the transformer. We demonstrate that the attention mechanism learns to select “good” sequence elements (those most relevant to the prediction task) while suppressing “bad” ones. This separation is formally framed as a convex optimization program, similar to classical support-vector machines (SVMs), but with a distinct operational interpretation. Ultimately, our findings enhance our understanding of the attention mechanism, enabling us to interpret it as a hierarchy of SVMs that effectively separate and select optimal sequence elements.

Refreshments served at 3:30pm 4th floor Lounge – Amos Eaton

Biographical Sketch

Davoud Ataee Tarzanagh is a post-doctoral researcher in the Department of Biostatistics, Epidemiology, and Informatics at the University of Pennsylvania, hosted by Li Shen and Qi Long. He is also affiliated with ASSET (Center for AI-enabled Systems: Safe, Explainable, and Trustworthy). Previously, he was a post-doctoral researcher in the Department of Electrical Engineering and Computer Science at the University of Michigan, hosted by Laura Balzano. He received his Ph.D. in Mathematics from the University of Florida in 2020, where he was co-advised by William W. Hager and George Michailidis. His research interests include mathematical optimization, high-dimensional data analysis, and trustworthy machine learning.

Davoud Ataee Tarzanagh
tarzanagh@gmail.com
University of Pennsylvania