THE UNIVERSITY OF CHICAGO

Computational and Applied Mathematics
&
Statistics Student Seminar

Adela DePavia

Computational and Applied Mathematics
University of Chicago

How do simple rotations affect Adam?

Wednesday, February 25, 2026
12:00 PM
Rosenwald 011
1115 E 58th St

ABSTRACT

Adaptive gradient methods—such as Adagrad, Adam, and their variants—have found widespread use in machine learning, signal processing, and many other settings. However, many algorithms in this family are not rotationally equivariant: in this talk we examine how a simple rotations in either parameter space or data space can drastically impact both the convergence rates and the generalization of these algorithms. We'll characterize how Adam's implicit bias can change under rotations of the training data distribution, and demonstrate that Adam's generalization can be extremely brittle to such rotations. We'll then describe a data-driven method proposed in our recent work which produces a "favorable" basis for adaptive algorithms, and demonstrate the potential of this reparameterization to make Adam's generalization more robust.