

NUMERICAL ANALYSIS AND SCIENTIFIC COMPUTING
SEMINAR

*Importance Sampling for Approximating High-Dimensional
Kernel Matrices*

Christopher Musco
New York University

Abstract: Randomized algorithms for compressing large matrices and accelerating linear algebraic computations have received significant attention in recent years. Unfortunately, popular methods based on Johnson-Lindenstrauss random projections and related techniques cannot be efficiently applied to dense implicit matrices that appear in many data applications – specifically they are difficult to apply to large kernel Gram matrices constructed from high-dimensional data. In this talk, I will discuss recent efforts to address this issue by instead compressing large kernel matrices using randomized importance sampling methods, and specifically those based on leverage scores. I will describe a recursive algorithm for leverage score sampling that provably computes a near-optimal rank- k approximation to any $n \times n$ kernel Gram matrix in roughly $O(nk^2)$ time, which is sublinear in the size of that matrix. Random projection methods on the other hand require at least $O(n^2)$ time. I will also touch on connections between leverage score sampling and the practically popular “random Fourier features method”, including insights on how to improve this method using tools from Randomized Numerical Linear Algebra.

Bio: Christopher Musco is an Assistant Professor at New York University’s Tandon School of Engineering. He received his Ph.D. from the theoretical computer science group at MIT in 2018. Christopher’s research focuses on scalable algorithms for matrix problems that arise in machine learning and data science, with a focus on randomized approximation algorithms.

Friday, February 19, 2021, 1:30 pm
<https://emory.zoom.us/j/95900585494>

MATHEMATICS
EMORY UNIVERSITY