March 6, 2023
2:00pm-3:00pm
KAP 414

Prof. Lingjiong Zhu
(Florida State University)

The Heavy-Tail Phenomenon in SGD

Abstract: Heavy-tail phenomena in stochastic gradient descent (SGD) have been reported in several empirical studies. Experimental evidence in previous works suggests a strong interplay between the heaviness of the tails and generalization behavior of SGD. To address this empirical phenomena theoretically, we establish novel links between the tail behavior and generalization properties of SGD through the lens of algorithmic stability. We show that the generalization error decreases as the tails become heavier, as long as the tails are lighter than a threshold. Moreover, we investigate the origins of the heavy tails in SGD. We show that even in a simple linear regression problem with independent and identically distributed data whose distribution has finite moments of all order, the iterates can converge to a stationary distribution that is heavy-tailed with infinite variance. We further characterize the behavior of the tails with respect to algorithm parameters, the dimension, and the curvature. We then translate our results into insights about the behavior of SGD in deep learning. We support our theory with experiments conducted on synthetic data, fully connected, and convolutional neural networks.

Zoom Link: USC Math Finance Colloquium
Time: Mar 6, 2023 02:00 PM Pacific Time (US and Canada)

Join Zoom Meeting
https://usc.zoom.us/j/97466624791?pwd=WU4wSGo4OXlYT2xUZzhBaUhWdlNRZz09

Meeting ID: 974 6662 4791
Passcode: 803148