UCLA

BIOSTATISTICS SEMINAR

SPRING 2016

Subgaussian estimators of the mean of a random matrix with heavy-tailed entries

Stanislav (Stas) Minsker , Ph.D.

Assistant Professor

Department of Mathematics

University of Southern California

Wednesday, May 11, 2016

3:30pm - 4:30pm, CHS 33-105A

Refreshments served at 3:00 PM in room 51-254 CHS

ABSTRACT:

In the past decade, the areas of high-dimensional statistics and statistical learning have seen numerous breakthroughs in structural estimation, concerned with a task of recovering a high-dimensional parameter that belongs to a set with ``simple'' structure from a small number of measurements.

Examples include high-dimensional sparse linear regression, low-rank matrix recovery and covariance estimation.

However, theoretical guarantees for popular techniques (e.g., Lasso and nuclear norm minimization) usually require strong assumptions on the underlying probability distribution, such as sub-gaussian noise.

What happens when these conditions are violated, which is the case in many applications? Can the assumptions be weakened without sacrificing the quality of theoretical guarantees?

Motivated by this question, we develop a new estimator of the (element-wise) mean of a random matrix, and show how it can be used to give a positive answer in some cases. Our estimator admits sub-gaussian or subexponential concentration around the unknown mean in the operator norm, assuming only two moments on the entries of the matrix.

We will explain the key ideas behind the construction, as well as applications to the matrix completion problem.