

November 10, 2014  
KAP 414  
2:00 PM – 3:00 PM

Stanislav Minsker  
Duke University and Wells Fargo Securities

***“Robust and scalable statistical estimation: a tale of the geometric median”***

**Abstract:** Contemporary high-dimensional data analysis problems pose several general challenges. One is related to resource limitations: massive data require computer clusters for storage and processing. Another problem occurs when available observations are contaminated by noise and "outliers" that are not easily identified and removed. An attempt to address these challenges raises natural question: can we design estimation techniques that

- (i) admit strong performance guarantees under weak assumptions on the noise and
- (ii) can be implemented in parallel while preserving the quality of estimation?

We make a step towards solving these problems and suggest a general method based on the properties of a *geometric median*, which is one of the possible extensions of a univariate median to higher dimensions.

Our approach will be illustrated through several examples: we will discuss tight estimators of the mean on a real line as a warm-up, and proceed to sparse linear regression and medians on the space of probability measures.