Introduction to High Dimensional Data Analysis
Math 465 - Fall 2014

Dr. Mauro Maggioni
Office: 319 Gross Hall
Phone: 660-2825
Web page: www.math.duke.edu/~mauro
E-mail: mauro.maggioni at duke

Course Synopsis

Prerequisites
Mathematics 103 (multivariable calculus) and Mathematics 104 (linear algebra) or equivalents and basic probability (at least discrete random variables) are needed for the mathematical background. Basic experience in programming in C or MATLAB or R.

Detailed course content

• High-dimensional data: problems and challenges. Density estimation, regression; curse of dimensionality.
• MATLAB tutorial
• Linear dimension reduction. Principal component analysis. Johnson-Lindenstrauss Lemma and random projections. Statistical and algorithmic
• Nonlinear dimension reduction. Graphs, operators on graphs, basics of spectral graph theory, eigenvalues and eigenvectors. Spectral clustering. Manifold learning, overview of various algorithms.
• Density estimation, in low- and high-dimensions. Fundamental results on kernel density estimation. Bias and variance; function spaces; minimax optimal rates of convergence.
• Regression, in low- and high-dimensions. Fundamental results in kernel regression.
• Connections to statistical signal processing: Fourier analysis and wavelets; dictionary learning.


Assignments
Weekly problem sets will include theory, analysis and computational projects. Requests for extensions on homework should be done before the due data; unexcused late assignments will be penalized. You are encouraged to discuss the homework problems with your classmates, but your final submission must be entirely your own independent work (see the Duke Community Standard).

Exams
Two midterms Grade to be based on weekly assignments (30%), two midterms (30%) and final project (50%).

Recommended
Some programming experience in Matlab or R; document preparation: \LaTeX; some more advanced probability (e.g. continuous random variables), some signal processing (e.g. Fourier transform, discrete and continuous); basic functional analysis.
Additional Information
Students from all areas of science, engineering, computer science, statistics, economics and quantitative studies that need advanced level skills in solving problems related to the analysis of data, signal processing, or statistical modeling are encouraged to enroll.