Optimal Learning

Peter Binev
University of South Carolina
binev@math.sc.edu

Learning an unknown function f from given data observations is a dominant theme in data science. The central problem is to use data observations of f to construct a function $\hat{f}$ which approximates f away from the data. There are numerous settings for this learning problem depending on (i) what additional information we have about f, (ii) how we measure the accuracy of how well $\hat{f}$ predicts f, (iii) what is known about the data and data sites. The main theme of this talk is twofold:

- to determine the optimal performance possible (the smallest possible error of recovery) in a given learning setting;
- to understand which discrete optimization formulations, when successfully numerically implemented, give a (near) optimal solution to the learning problem.

The remaining step is then to give a viable numerical method with convergence guarantees and bounds on computation which solves the discrete optimization formulation. It often remains an open question as to whether a proposed numerical optimization strategy, such as gradient descent methods, actually converges in the given learning setting to a near optimal solution.

This talk is concerned with evaluating how well an approximation $\hat{f}$ performs and determining the best possible performance among all choices of an $\hat{f}$. Given answers to these fundamental questions, one can then turn to the construction of numerical procedures and evaluate their performance against the known best possible performance.

Joint work with: Andrea Bonito, Ronald DeVore, and Guergana Petrova.