Math 89S   Spring 2013

Instructor: Paul Bendich   **Office:** 210 Physics Building
Phone: 660-2811   **e-mail:** bendich@math.duke.edu

Title: The Emerging Science of Complex Data

Description: Over the last decade or so, researchers in areas as diverse as astronomy, environmental science, image analysis, biology, high-energy physics, psychology, and finance, have been confronted with a very welcome but still vexing problem. As computing power has grown, they suddenly have large amounts of data which is often highly complicated, poorly understood, and rapidly growing. In order to make further progress, these researchers must find ways to analyze this data, extract interesting structural features from it, and reduce its size without destroying important content.

This course will introduce students to some of the potential achievements and pitfalls of this new scientific challenge, as well as to the mathematical and statistical methods currently being developed and used to analyze and tame these large datasets. Some of these methods make use of linear algebra, geometry and topology, and so students will be exposed to interesting applications of these subjects without necessarily having to take formal courses in them.

Many of the methods are also at present inadequate in many ways, and so students will have the opportunity to understand the current limitations of mathematics and discuss ways to make improvements. We will discuss the ways in which the need to analyze large datasets has led to the formation of new collaborative paradigms in science. We will also consider some of the current and past uses and abuses of data analysis in social policy.

In addition to background reading, discussions, and problem sets, students will write a research paper: students may identify a particular dataset of broad interest, and survey the current methods used to analyze it, along with their limitations, or students may delve into a particular data analysis technique, exploring its applications and limitations. After spring break, students will present their findings in class, and will then be expected to revise and rewrite their papers, taking into account the critiques of their classmates.

This seminar is offered as part of a Research Training Grant which involves, among other things, a summer research program. Highly motivated seminar students will be

invited to participate in this program, exploring new problems in small groups with more senior students.