Course Announcement Spring 2010

M 532 Mathematical Modeling of Large Data Sets

Description: A time-honored approach for the investigation of unexplained phenomena is to attempt to infer laws, or explain processes, from the patterns present in collected data. However, our phenomenal ability to acquire data has outstripped our ability to process and analyze it. Thus, researchers today are confronted with a modern dilemma. Presumably the more information available concerning a phenomenon the better. Yet, a massive data set storing the information, in and of itself, a potentially significant barrier to the investigation. We will present an overview of several mathematical tools for overcoming problems associated with analyzing high-dimensional and massive data sets. Our approach is geometric in nature and the main tool is the dimensionality reducing mapping. These mappings are required for the analysis and representation of information (patterns) in large data sets generated by physical or numerical experiments. The main techniques to be presented include optimal orthogonal expansions, Fourier analysis, radial basis functions, neural networks and wavelets. These approaches are developed with a goal towards understanding the geometric aspects of data processing. We will emphasize the mathematical similarities and differences of the procedures in the context of

- data reduction of massive data sets
- modeling of empirically observed phenomena
- signal separation and missing data
- nonlinear signal and image processing
- dynamical systems modeling
- data on manifolds, algebraic sets

Prerequisites: The class assumes no prior knowledge of the field of pattern analysis. It is expected that the student will have a mathematics background typical of a graduate student in applied mathematics or engineering including linear algebra, e.g., M369 (M560 preferred). Knowledge of some programming language, e.g. Matlab, Fortran or C++ is also required. Homework: This will consist of computer assignments as well as the mathematical foundations of the subject. There will be approximately six problem sets and one project all carrying equal weight in computing the final grade. Background Text: Geometric Data Analysis, Michael Kirby, Wiley & Sons Time and Place: t.b.a. Information: Professor: Michael Kirby, Phone: 491-6850, Email: kirby@math . colostate