1 MATH892 Methods in Computational Mathematics

Model Reduction and Information Geometry

Description and goals. Constructing concise descriptions is a central objective of the physical, biological, or social sciences. Traditionally, this has been accomplished by researchers using particular insight in the phenomena under study to construct suitable models. Recent advances in computational techniques provide an additional set of tools that seek to integrate the large amount of data generated by simulations, sensors or surveys.

This course provides an introduction to the mathematical theory of model reduction, i.e., generally applicable techniques to extract essential features of a phenomenon from available simulation and measurement data. The course focus is on practical implementation of the theory with a strong computational component. This material often receives the moniker: "Data-driven discovery".

At the end of each course module, students should have attained a basic familiarity with the practical application of the mathematical theory to a problem of practical interest chosen from the student's research objectives.

Organization. The course consists of three modules (5 weeks each), of increasing abstraction:

1. *Model reduction based on linear projection in Euclidean spaces*. Linear approximation of systems by the methods of statistics or numerical simulation of ODEs, PDEs, integral equations. Linear projection using various norms.

2. *Information theoretic model reduction*. Introduction to information theory and information functionals (entropy, Bregman divergences, Kullback-Leibler divergence). Model reduction based on approximation of information functionals.

3. *Information geometry model reduction*. Introduction to Riemannian differential geometry of manifolds of statistical descriptions of systems. Nonlinear projection by transport along geodesics of underlying manifold. Riemannian differential geometry will be introduced as necessary.

Prerequisites. Basic coding ability is assumed in both a compiled (e.g. C/C++, Fortran) and interpreted (Matlab/Octave, Python) language. Familiarity with a symbolic computational package (Mathematica/Maple/Macsyma) will be highly useful. The course will use the following topics from various mathematical disciplines.

- *Computational linear algebra* (normed spaces of vectors and matrices, projection, leat-squares problems, QR, LU, SVD, Cholesky, eigenvector matrix factorizations with associated algorithms).

- *Vector calculus*. Gradient, divergence, curl, associated integral theorems (e.g., Stokes, Green formulas). Taylor series in $\mathbb{R}^n$.

- *Differential geometry of curves and surfaces*. Curves and surfaces in $\mathbb{R}^3$. Tangent, normal vectors, Frenet triad, curvature, Gauss curvature forms.

- *Numerical analysis*. Approximation of univariate and multivariate functions, numerical integration and differentiation, analysis of approximation convergence.

Bibliography. Course lecture notes will be provided. Consultation of the following material will be useful:

- *Approximation of Large-Scale Dynamical Systems*, A.C. Antoulas, SIAM, 2005

- *Elements of Information Theory*, T.M. Cover and J.A. Thomas, Wiley, 1991.
- *Probability Theory: The Logic of Science*, E.T. Jaynes, Cambridge, 2003.
- *Methods of Information Geometry*, S.-I. Amari, AMS, 1993
- *Matrix Information Geometry*, F. Nielsen, R. Bhatia, Springer, 2013.

Coursework and grading. Coursework consists of suggested exercises (self-graded against posted solutions) and projects for each course module. Each project applies mathematical methods discussed in the corresponding module to a problem relevant to student's research.