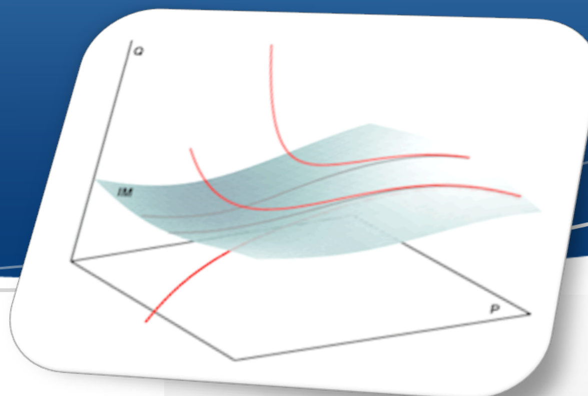
Applied Math Seminar

Spring 2015

United States Naval Academy
Tuesday March 24, 2015
12 noon, Room CH 351



The Geometry of Data

Many representations of text, audio, image and video data are high-dimensional point clouds. The curse of dimensionality is the observation that as data is spread into high dimensions, the distance between points becomes large and the corresponding density very low and difficult to estimate. In order to avoid this issue, one imagines that only the data representation is high dimensional but the data actually lies along curved low-dimensional structures characteristic of a generative mechanism with limited degrees of freedom or invariance properties related to symmetries. The manifold hypothesis proposes that these structures may have continuity properties that allow their representation as a Riemannian manifold with the local properties of Euclidean vector space and a smooth (differentiable) connection.

To learn a manifold from point cloud data is challenging and requires a mathematical framework to evaluate the continuity of tangent spaces at different points as determined by a probability model. At the intersection of statistics, Riemannian geometry and information theory lies the topic of information geometry. This approach uses an information-theoretic metric to determine the geodesic distance between probability distributions on the curved manifold defined by its statistical parameters. The manifold of multivariate Gaussian distributions is a particularly important example and has a non-Euclidean geometry that demonstrates the connection between statistical estimation and spaces with negative curvature. Solutions to the Laplace-Beltrami operator on this manifold form the basis for constructing a useful integral transform. Expansions over the eigenfunctions of this operator in Euclidean space constitute the Fourier transform and in spaces of constant positive curvature, the spherical harmonics. Mathematical speculation is offered for a new type of integral transform and the application of the Ricci flow, which played an important role in solving the Poincare Conjecture, to manifold learning.

Speaker

Dr. Jeff M. Byers
jeff.byers@nrl.navy.mil
Materials and Sensors Branch

Naval Research Laboratory, Washington, DC

Organizer:
Reza Malek-Madani, United States Naval Academy rmm@usna.edu