Testing the Manifold Hypothesis
Hariharan Narayanan, Massachusetts Institute of Technology

Tuesday, March 29, 2011 - 2:30pm

Increasingly, we are confronted with very high dimensional data sets. As a result, methods of avoiding the curse of dimensionality have come to the forefront of machine learning research. One approach, which relies on exploiting the geometry of the data, has evolved into a subfield called manifold learning.

The underlying hypothesis of this field is that due to constraints that limit the degrees of freedom of the generative process, data tend to lie near a low dimensional submanifold. This has been empirically observed to be the case, for example, in speech and video data. Although there are many widely used algorithms motivated by this hypothesis, the basic question of testing this hypothesis is poorly understood. We will describe an approach to test this hypothesis from random data.

Related Links:
Pacific Institute for the Mathematical Sciences