Bocconi
Università Commerciale Luigi Bocconi

Department of Decision Sciences

Statistics Seminar

Looking for manifolds and ridges

Isabella Verdinelli
Università di Roma, La Sapienza

Thursday, 3rd October 2013
12:45pm Room 3-E4-SR03 Via Rontgen 1 Milano

Abstract

This talk presents some nonparametric methods for estimating stable, high-density regions in point clouds.
I will first discuss the formal problem of locating hidden structures in D-dimensional noisy data. I will concentrate on the general problem of estimating a d(< D)-dimension manifold that can be either hidden in the data or that has generated the data set. I will present the statistical model and show how the difficulty of this problem can be formalized using statistical minimax theory. I will specify the different minimax rates for various models for the noise, and brefly schetch some proofs.
In particular, for the case of Gaussian noise, the rate is extremely slow (logarithmic). A possible way out is estimating a surrogate for the manifold. This is a set that is close to the manifold and can be estimated at a polynomial rate.
A particular surrogate for a manifold is denoted by hyper-ridge (or ridge). A ridge is a d-dimensional set characterized by conditions on the eigenvalues of the Hessian. Ridges are shown to be estimated at polynomial rate and thus serve as a good surrogate for the underlying manifold.
In fact, the methods presented work well even when the underlying structure is not a manifold. I will also describe the problem of ``dimensional leakage'', in which structures can leave their imprints in several different dimensions.
Finally, I will comment about how the bootstrap can be used to assess the variability of the procedure.

Department of Decision Sciences

Via Röntgen 1 - 20136Milano

Tel. 02 5836.5632
Fax 02 5836.5630