STANFORD UNIVERSITY

DEPARTMENT OF STATISTICS

DEPARTMENT SEMINAR

4:15 p.m., Tuesday, March 12, 2002
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Micheal Jordan
UC Berkeley

Kernel Independent Component Analysis

Independent component analysis (ICA) refers to a class of latent variable models in which a set of "sources" must be recovered from an unknown set of linear mixtures of these sources. The problem is reminiscent of factor analysis, but whereas factor analysis assumes Gaussian latent variables, ICA is meaningful only for non-Gaussian latent variables; moreover, the distribution of the latent variables is assumed to be unknown. The problem is thus a semiparametric problem—we wish to estimate the mixing matrix under any (non-Gaussian) distribution for the sources.

We present a new class of algorithms for ICA which use contrast functions based on canonical correlations in a reproducing kernel Hilbert space (RKHS). On the one hand, we show that our contrast functions are related to mutual information and have desirable mathematical properties as measures of statistical dependence. On the other hand, building on recent developments in RKHS methods, we show that these criteria and their derivatives can be computed efficiently. Minimizing these criteria leads to flexible and robust algorithms for ICA. We illustrate with simulations involving a wide variety of source distributions, showing that our algorithms outperform presently known algorithms.

[Joint work with Francis Bach].