2nd International Symposium on
Information Geometry and its Applications
December 12-16, 2005, Tokyo

INFORMATION GEOMETRY OF BREGMAN DIVERGENCE AND ITS APPLICATION TO LEARNING ALGORITHMS

NOBORU MURATA

Waseda University

In information geometry, discrepancy measures between distributions play a key role, and there are many candidates for such measures. In this talk, we discuss geometrical properties of the Bregman divergence, which is a class of divergences derived from convex functions and includes KLdivergence as a special case.

The Bregman divergence has some interesting properties when it is applied to statistical inferences. For example, estimators derived from the Bregman divergence is not Fisher efficient in general, but it shows robustness to outliers and noises (Minami & Eguchi, 2002; Takenouchi & Eguchi, 2004; Fujisawa & Eguchi, 2005 and so on). Another interesting property is its duality. The Bregman divergence is associated with two representations of distributions, one is mixture representation, and the other is representation based on the convex function. Like the R-divergence, these two representations are related to a dualistic property, and one of the characteristics is simply expressed by the Pythagorean relation.

In this talk, we focus on two kinds of algorithms: hEM algorithmh which is used for estimating probability densities of mixture models, and hboosting algorithmh which is used for constructing complex discriminant functions by combining simple ones. We discuss their properties from geometrical viewpoints based on the duality of the Bregman divergence.