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

STATISTICAL CURVATURE AND STOCHASTIC COMPLEXITY

JUNICHI TAKEUCHI

We discuss the relationship between the statistical embedding curvature and the logarithmic regret (regret for short) of the Bayesian prediction strategy (or coding strategy) for curved exponential families and Markov models. The regret of a strategy is defined as the difference of the logarithmic loss (code length) incurred by the strategy and that of the best strategy for each data sequence among a considered class of prediction strategies [Shtarkov '88]. (The considered class is also called a reference class.) Since a prediction strategy is equivalent to a probability distribution, the class of prediction strategy is equivalent to a statistical model. Note that the logarithmic loss (equivalent to code length) by the minimax strategy is equals to Rissanen's stochastic complexity.

For this matter, it can be shown that the Bayesian strategy with Jeffreys prior (Jeffreys strategy for short) asymptotically achieves the stochastic complexity upto the constant term, when the considered class is an exponential family. This is due to the fact that the logarithmic loss of Bayes mixture strategy depends on the exponential curvature of the considered class. Hence, it can be shown that the Jeffreys strategy is asymptotically minimax, if and only if the considered class is an exponential family. That is, the Jeffreys strategy asymptotically achieves the SC, if and only if the exponential curvature equals 0.

We also consider the expected version of regret (known as redundancy in information theory field). When the actual probability distribution belongs to the reference class, the Jeffreys strategy asymptotically achieves the minimax redundancy, irrelevant to the curvature of the reference class as shown by Clark and Barron (1994). However, if the actual probability distribution does not belong to the reference class, the situation differs. In that case, the redundancy of Jeffreys strategy depends on both exponential curvature and mixture curvature of the reference class.