Minimax Regret Rates for Prediction under Log Loss via Information Geometry

Nir Weinberger
The Viterbi Faculty of Electrical and Computer Engineering
Technion - Israel Institute of Technology
nirwein@technion.ac.il

Meir Feder
School of Electrical Engineering
Tel Aviv University
meir@eng.tau.ac.il

Abstract--- We consider the problem of learning a predictive model from given data samples in which the predictor's quality is measured by the log loss. We focus on the misspecified setting, in which the true predictive model generating the data is chosen from a set different from the possible models that can be chosen by the learner. We establish minimax regret upper and lower bounds in term of projected covering and packing entropies, derived from reverse information projections of the data predictive models on the learner predictive models. We also consider the setting in which the features density is unknown to the learner. We characterize the increased regret in terms of covering entropies in the space of feature densities for a metric chosen judiciously according to the sensitivity of the predictive model to the features.