論文の内容の要旨

論文題目: Theory of Information Integration in Statistical Learning
（統計的学習における情報統合の理論）

氏名　鈴木　大慈

This thesis theoretically investigates several kinds of information integration in statistical learning. In the statistics and statistical learning literature, information integration is realized through various types of methods. For example, Bayesian prediction integrates distributions over the model, Boosting integrates weak classifiers or functions to obtain strong ones, and multi-task learning integrates data and shares information around relevant learning tasks. We address six main topics of information integration. The first one is to analyze a prior distribution selection problem of a general class of Bayesian predictive distribution. We see that asymptotic risk is characterized by a second order elliptic operator which is a generalization of the Laplacian operator. This says that positive harmonic function improves the predictive performance. The second contribution is to propose a new Boosting method which combines Bayesian estimation and Boosting, show experimentally the effectiveness of our proposed method, and give a consistency proof of the proposed algorithm. The third one is to analyze data integration over classification tasks parameterized by a continuous variable. Integrating data around relevant tasks is often done to improve the performance in multi-tasks learning setting. We analyze it by assuming continuity between tasks, and the optimal range from which data are integrated and the optimal learning rate are given. We show that this learning rate achieves the mini-max lower bound which cannot be improved by any learning protocol. The fourth one is to give a parametric analysis of data integration as dealt with in the third one. The asymptotic risk is given by the Laplacian operator, thus geometric meaning of the risk is clarified. The fifth one is to analyze data integration on a discrete set of classification tasks. We observe that data integration could improve the generalization performance. The Final one is to analyze a density ratio estimator which is applied to covariate shift setting where input distributions for training and test are different. We give convergence rate on non-parametric settings and a variant of asymptotic normality on parametric settings. We also propose a new mutual information estimation method utilizing

density ration estimation. Our proposed algorithm has several attractive properties and numerical experiments support the usefulness of our proposed method. Theoretical justification is also given. This thesis gives several fundamental tools for information integration studies.