Research Plan

In my future research, I plan to continue working on important open research questions that I have addressed so far, but I will focus mainly on the research topic "Evaluation of Information Criteria Based on High-Dimensional and Large-Sample Asymptotic Theory". One of the important properties of information criteria for model selection is consistency, which is the property that the probability of selecting the true model asymptotically approaches 1. Under the large-sample asymptotic theory so that only the sample size goes to infinity, known as the ordinary asymptotic theory, it is known that the AIC do not have consistency, while the Bayesian information criterion (BIC) has consistency. On the other hand, in multivariate data, when the number of dimensions of response variable vectors is large compared to the sample size, it is known that the approximation accuracy of probability distributions becomes poor in asymptotic approximations based on the large-sample asymptotic theory. In the case of high-dimensional data, where the dimension is large, it is known that if the asymptotic approximation derived from the high-dimensional and large-sample asymptotic theory so that the sample size and the dimension go to infinity simultaneously is used, the approximation accuracy of the distribution does not become poor. Fujikoshi, Sakurai, and Yanagihara (2014) and Yanagihara, Wakaki, and Fujikoshi (2015) obtained results contrary to the well-known result that in the variable selection problem in multivariate linear regression models, when consistency is evaluated using high-dimensional and large-sample asymptotic theory, the AIC may be consistent while the BIC may not be consistent. In particular, Yanagihara, Wakaki, and Fujikoshi (2015) derive conditions for the penalty term to satisfy consistency when the penalty term of the information criterion is generalized. However, under these conditions, whether or not consistency is achieved depends on the divergence rate of the maximum eigenvalue of the non-centrality parameter matrix. Unfortunately, no existing criteria could be found to ensure consistency for all non-centrality parameter matrices. Therefore, Yanagihara *et al.* (2017), Yanagihara (2019), and Oda *et al.* (2020), and Oda & Yanagihara (2020) proposed consistent criteria for any non-centrality parameter matrix, regardless of whether large-sample asymptotic theory or high-dimensional and large-sample asymptotic theory are used for evaluating consistency, by reevaluating the conditions for consistency. In particular, it is noteworthy that Yanagihara (2019) succeeds in removing the true model normality assumption assumed in other results. Proposing such criteria in multivariate models other than multivariate linear regression models will be a major challenge in the future. If such criteria can be proposed, it will be possible to use information criteria that guarantee consistency regardless of the size of the dimensions of response variable vectors and the divergence speeds of non-centrality parameter matrices, which are usually unknown, and to propose user-friendly methods for analysts. Furthermore, in model selection, in addition to consistency, the asymptotic loss efficiency, the property that the ratio of the loss function of the selected model to the minimum value of the loss function converges to 1, and the asymptotic average efficiency, the property that the ratio of the expected value of the loss function of the selected model to the minimum value of the expected value of the loss function converges to 1, become important. In particular, if variable selection is performed using an efficient criterion, it can be said that this is a good variable selection method in terms of asymptotic prediction. Therefore, I attempt to propose an efficient information criterion for any non-centrality parameter matrix by reevaluating the conditions for efficiency. Based on these results, I believe that user-friendly methods can be proposed to analysts to improve prediction accuracy.

Although it is not directly related to my research, I would like to continue providing statistical consultation services that I was involved in at my previous institution. While my previous institution was a university, most of the statistical analysis consultations I received were related to master's and doctoral theses of graduate students, as well as statistical analysis methods for faculty research.

However, I also received consulting requests from several companies. As I mentioned before, statistical consultations do not necessarily lead to large-scale joint research projects right away, but I believe that these kinds of steady activities will serve as a stepping stone for future large-scale joint research projects.