Japanese Journal of Statistics and Data Science
https://doi.org/10.1007/s42081-024-00250-5

PREFACE

Special feature: Stein estimation and statistical shrinkage methods

Yuzo Maruyama[1]

The twelfth special issue of the Japanese Journal of Statistics and Data Science (JJSD) focuses on statistics for shrinkage estimation. In recent years, statistical shrinkage methods have become increasingly prevalent for addressing scientific inquiries, particularly in experiments and scenarios involving inference and prediction based on large-scale data. Initially, the effectiveness of Stein-type shrinkage estimation relied on statistical decision theory and empirical Bayes inference. Present-day methodologies in shrinkage methods offer scalable inference techniques to tackle intricate problems across various domains, including the biological sciences, signal processing, economics, and finance. Despite the flourishing development of modern shrinkage methods, numerous theoretical and computational challenges persist. This special feature comprises fourteen papers aimed at showcasing recent research and reviews pertinent to Stein estimation and statistical shrinkage methods. Within these papers, readers will encounter a diverse array of theories, methodologies, and empirical analyses. The papers in the special feature are summarized as follows:

The first article is "Machine learning and the James–Stein estimator" by Bradley Efron. The seminal work of James and Stein on estimating a multivariate normal mean vector made a spectacular first impression on the statistical community through its demonstration of inadmissibility of the maximum likelihood estimator. Efron points out that it continues to be influential, but not for the initial reasons. In the 1970s, Efron and Carl Morris collaborated on a series of papers where they justified the James–Stein estimator as an empirical Bayes shrinkage procedure. Furthermore, as Efron highlights, modern topics such as Benjamini and Hochberg's false discovery rate algorithm can trace their roots back to the empirical Bayes method. Tragically, Morris passed away during the publication process. In tribute to Carl Morris, Efron dedicates this article to his memory.

The second article is titled "Stein's identities and the related topics: an instructive explanation on shrinkage, characterization, normal approximation and goodness-of-fit" by Tatsuya Kubokawa. Stein-type identities are renowned for their significant

✉ Yuzo Maruyama
maruyama@port.kobe-u.ac.jp

1 Graduate School of Business Administration, Kobe University, 2-1 Rokkodai, Nada, Kobe 657-8501, Japan

Published online: 11 May 2024

utility and effectiveness in deriving shrinkage estimators that enhance crude estimators across normal, gamma, Poisson, and negative binomial distributions. This review article provides a thorough and illuminating explanation of the applications of Stein-type identities. It is worth noting that Stein's identities are effectively utilized in most papers within this volume.

The third article "Matrix quadratic risk of orthogonally invariant estimators for a normal mean matrix" by Takeru Matsuda deals with estimation of a normal mean matrix under the matrix quadratic loss. Utilizing the Stein-type identity, he puts forward a universal formula for the matrix quadratic risk of orthogonally invariant estimators. This formula enables the calculation of the matrix quadratic risk of a singular value shrinkage estimator, inspired by Stein's suggestion for enhancing the Efron-Morris estimator.

The estimation of the multivariate normal mean under the quadratic loss function is primarily addressed in the fourth through seventh articles. The fourth article, titled "On priors which give Bayes minimax estimators of Baranchik's form" authored by Fourdrinier, Strawderman, and Wells, focuses on the minimaxity of shrinkage estimators, often assured by Baranchik's condition in this field. They note that most well-known Bayes minimax estimators have been suggested under the prior distribution using scale mixtures of normal distributions. However, they discover certain priors that cannot be depicted as scale mixtures of normals but still result in Baranchik-type minimax estimators.

The fifth article is titled "Shrinkage estimation with logarithmic penalties" by Tatsuya Kubokawa. His approach relies on the equation of the first-order condition with a logarithmic penalty, introducing both one-step and two-step shrinkage estimators. The one-step estimator closely resembles the James–Stein estimator, while the differentiable two-step estimator shows similar performance to the positive-part Stein estimator. Despite the latter not meeting Baranchik's conditions, both estimators can be proven to be minimax.

The sixth article is titled "Non-minimaxity of debiased shrinkage estimators" by Maruyama and Takemura. The debiased shrinkage estimator, designed to shrink towards the origin for smaller values and exactly match the unbiased estimator for larger values, appears to be minimax. However, they demonstrate that it does not hold the minimax property under mild conditions.

The seventh article is titled "Expansion estimators improving the bias and risk of James–Stein's shrinkage estimator" by Hisayuki Tsukuma. Firstly, he examines bias correction of the James–Stein shrinkage estimator through expansion. Subsequently, he demonstrates that certain expansion estimators simultaneously reduce the bias and improve the risk of the James–Stein estimator.

The eighth paper is titled "Asymptotics of coverages of HD confidence sets and recentering at shrinkage estimates: phase transitions, large deviations" by Anirban DasGupta. He discusses confidence set estimation of the mean vector of a p-dimensional normal distribution, utilizing n independent and identically distributed (iid) observations from it. Through a series of theorems, he presents the coverage properties, expected volume, and girth of both the classic Hotelling confidence set and Bayesian Highest Posterior Density (HPD) sets under normal priors. Notably, in the scenario where p increases with n, he demonstrates an explicit phase transition in the coverage probability of the Hotelling confidence set and provided subexponential bounds on the noncoverage probability of HPD sets.

The utility of the shrinkage procedure under the Poisson model is outlined in the ninth and tenth articles. The ninth article is titled “Bayesian prediction and estimation based on a shrinkage prior for a Poisson regression model” by Fumiyasu Komaki. He introduces a Poisson regression model with a natural structure, approached from the standpoint of information geometry. Within this framework, he shows that the Bayes extended estimator and Bayesian predictive density, derived from the shrinkage prior, outperform those based on the Jeffreys prior when considering the Kullback-Leibler loss.

The tenth article is titled “Bayesian shrinkage estimation for stratified count data” by Yasuyuki Hamura. He delves into simultaneous estimation methods for count data scenarios where both direct and aggregated observations are accessible. While these situations are common in practice, generalizing results from the literature is not always straightforward, unlike when dealing with continuous data using the normal distribution. In this article, he establishes conditions for dominance and subsequently proves minimaxity and admissibility of the estimators he proposes.

The eleventh article is titled “An adaptive singular value shrinkage for estimation problem of low-rank matrix mean with unknown covariance matrix” by Yoshihiko Konno. For the problem of estimating a covariance low-rank matrix, he considers estimators with soft-thresholding singular values and proposes a formula to choose a threshold based on the SURE method.

The twelfth article is titled “On combining unbiased and possibly biased correlated estimators” by Zinonos and Strawderman. They study estimators that combine an unbiased estimator with a possibly biased correlated estimator of a mean vector. These combine estimators adopt a shrinkage-type approach, wherein the unbiased estimator is shrunk towards the biased estimator. They introduce conditions under which the combined estimator outperforms the original unbiased estimator.

The thirteenth article is titled “Stein-rule M-estimation in sparse partially linear models” by Raheem, Ahmed and Liu. They introduce and analyze the statistical properties of shrinkage M-estimators using Stein-rule estimation for partially linear models, assuming sparsity. They examine the asymptotic properties of these shrinkage estimators through both analytical and numerical methods. Additionally, they derive the asymptotic bias and risk of the estimators in closed form.

The final article of this volume is titled “Confidence interval for normal means in meta-analysis based on a pretest estimator” by Taketomi, Chang, and Emura. Meta-analysis, a statistical technique used to consolidate quantitative findings from a collection of published studies, has seen recent advancements. A frequentist estimator was introduced for individual study means, employing the pretest (preliminary test) estimator. However, the construction of confidence intervals for the pretest estimator in meta-analysis has not been addressed. They propose an innovative method to establish confidence intervals based on the pretest estimator for meta-analysis.

Finally, I would like to extend my sincere gratitude to the former Editor-in-Chief, Professor Makoto Aoshima, and the current Editor-in-Chief, Professor Hiroki Masuda, for granting me the opportunity to oversee this special feature. I also want to express my gratitude to all the authors who have contributed their engaging original works to this special feature. I am deeply appreciative of all the reviewers for their invaluable assistance in the peer-review process.