Bootstrapping the Stein-Rule Estimators

難波, 明生

Noname manuscript No.
(will be inserted by the editor)

Bootstrapping the Stein-Rule Estimators

Akio Namba

the date of receipt and acceptance should be inserted later

Abstract In this paper we consider the Stein-rule estimator and the positive-part Stein-rule estimator for the mean of a multivariate normal distribution and analyze the validity of the bootstrap methods for these estimators. We show that the conventional bootstrap is not always consistent and propose an alternative bootstrap method which is consistent when the conventional bootstrap is inconsistent. We also show the consistency of the *m* out of *n* bootstrap. Moreover, we propose an consistent bootstrap method based on a pre-test. Our simulation results show the validity of the proposed bootstrap in various setups.

Keywords Stein-rule Estimator, *m* out of *n* bootstrap, Centered bootstrap, Pre-test bootstrap
JEL Classification C13, C18

Declarations
Funding This work was supported by JSPS KAKENHI Grant Number 18K01546.
Conflicts of interest/Competing interests Not applicable.
Availability of data and material Not applicable.
Code availability FORTRAN code is available from the author on request.
Ethics approval Not applicable.
Consent to participate The author consents to participate in the research.
Consent for publication The author consents to publish the manuscript in Journal of Quantitative Economics.

Acknowledgment

The author is grateful to the anonymous referees and editors for kind and helpful comments and suggestions which greatly improves the analyses and the presentation of the manuscript. This work was supported by JSPS KAKENHI Grant Number 18K01546.

Graduate School of Economics
Kobe University
Rokko, Nada-ku, Kobe 657-8501
Japan
E-mail: namba@econ.kobe-u.ac.jp

1 Introduction

In estimating the mean of a multivariate normal distribution, the sample mean is the maximum likelihood estimator (MLE) and the best asymptotically normal (BAN) estimator. However, in small samples, Stein (1956) and James and Stein (1961) showed that the Stein-rule (SR) estimator has smaller mean squared error (MSE) when the dimension of the normal distribution is larger than or equal to three. Also, as is shown in Baranchik (1970), though the SR estimator has smaller MSE than the MLE, the positive-part Stein-rule (PSR) estimator has smaller MSE than the SR estimator. After these seminal contributions an extensive literature on the variants of the SR estimator and their performances has been developed. Those estimators are called the shrinkage estimators because they are usually obtained by shrinking the MLE to the origin. Recently, shrinkage estimators are applied to several general models [see e.g., Hansen (2016a, 2017)]. Also, it is well known some recent techniques including LASSO, SCAD and model averaging can be considered as kinds of shrinkage estimation [see, e.g., Fan and Li (2001), Hansen (2016b), Tibshirani (1996), Ullah et al. (2016) and Zou (2006)]. In particular, extending the idea of Hausman (1978) pre-test in Guggenberger (2010), Hansen (2017) proposed an estimator which consists of the ordinary least squares estimator and the two-stage least squares estimator under the assumption that the correlation coefficient is local to zero. He showed that the proposed estimator has an asymptotic structure similar to the Stein-rule estimator and that the obtained estimator has remarkable performance. Further, as is discussed in Ohtani (1993), the PSR estimator can be considered as a pre-test estimator. These results show that the ideas of shrinkage and pre-test estimation can be useful in applied analyses.

However, one of the problems concerning the application of the shrinkage estimators is the difficulty in assessing the accuracy of the estimators because the distributions of the shrinkage estimators are complex. For example, Ullah (1982) derived an approximate distribution function of the SR estimator. Also, Phillips (1984) derived the exact density function. Both of their results show that the distribution of the SR estimator is complex and it depends on the unknown parameters. Since the MSE of the shrinkage estimators are smaller than the MSE of the MLE, we can infer that the distributions of the shrinkage estimators concentrate around the mean more than that of the MLE. This means we may be able to obtain confidence regions which have smaller volume than the conventional MLE-based confidence region if we can obtain approximate distributions of the shrinkage estimators and utilize them. However, because of the complexities of their distributions, it is difficult to utilize the SR estimator or its variants in construction of confidence regions or hypothesis tests.

When the distribution of a statistic is unknown or complex, the bootstrap methods proposed by Efron (1979) may be applied. Some authors applied the bootstrap methods to the shrinkage estimators and examined their validity. For example, Brownstone (1990), Yi (1991) and Kazimi and Brownstone (1999) showed the validity of the bootstrap application to the SR estimator by simulations. Also, Vinod (1995) applied the double bootstrap to the ridge regression estimator, and showed the validity by simulations. On the other hand, Beran (1997) theoretically showed that the bootstrap approximation is inconsistent for the SR estimator and proposed a consistent parametric bootstrap method based on a pre-test for unknown parameters. Also, Wei et al. (2016) stated that the conventional bootstrap is inconsistent for shrinkage estimators, and proposed to apply the m out of n bootstrap to approximate the distribution of the shrinkage estimators. As will be shown in section 3, the conventional bootstrap is inconsistent when the mean of the distribution is 0, however, Wei et al. (2016) showed that the m out of n bootstrap is consistent even when the mean of the distribution approaches to 0, more precisely, the mean is $o(n^{-1/2})$ where n is the sample size.

They also proposed a way to obtain an optimal sample size m in the bootstrap resampling. From these results, though Beran (1997) showed the inconsistency of the conventional bootstrap for the shrinkage estimators, it seems that some bootstrap methods yield reasonable results when they are applied appropriately.

Thus, in this paper, we consider the SR estimator and the PSR estimator, and analyze the validity of the bootstrap methods. The organization of the paper is as follows. In section 2, we introduce the SR and the PSR estimators. In section 3, we theoretically analyze the validity of the bootstrap methods for the SR and the PSR estimators. It is shown that the conventional bootstrap for the SR and the PSR estimators is inconsistent at a point, i.e., the mean vector is 0. Also, we propose the centered bootstrap which is consistent at the point where the conventional bootstrap is inconsistent. Further, we show the consistency of the m out of n bootstrap. We also proposed a consistent nonparametric bootstrap method based on a pre-test in a similar way to Beran's (1997) parametric bootstrap. In section 4, we execute Monte Carlo simulations to investigate the performances of the methods introduced in section 3. Finally some concluding remarks are given in section 5.

2 The SR estimators

Assume that the $k \times 1$ vector $X_i \sim \mathrm{N}(\mu, \mathrm{I}_k)$, $i = 1, 2, \ldots, n$. Then, the ML estimator for the $k \times 1$ mean vector μ is given by $\widehat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. It is well known that the ML estimator is the BAN estimator and its covariance matrix can be obtained from Fisher's information matrix.

However, as shown in Stein (1956) and James and Stein (1961), the SR estimator given by

$$\widehat{\mu}_{\mathrm{SR}} = \left(1 - \frac{k-2}{n\overline{X}'\overline{X}}\right)\overline{X}, \tag{1}$$

has smaller MSE than the ML estimator when $k \geqslant 3$. Also, as is shown in Baranchik (1970), the PSR estimator given by

$$\widehat{\mu}_{\mathrm{PSR}} = \max\left[0,\ 1 - \frac{k-2}{n\overline{X}'\overline{X}}\right]\overline{X}, \tag{2}$$

further dominates the SR estimator in terms of MSE.

Since the SR and the PSR estimators have smaller MSE than the ML estimator, the distributions of the SR and the PSR estimators concentrate around the mean more than the distribution of the ML estimator. As for the small sample distribution of the SR estimator, Ullah (1982) derived an approximate distribution function of the SR estimator, and Phillips (1984) derived the exact density function. Their results show that the distribution of the SR estimator depends on unknown parameters. Therefore, it is difficult to obtain a small sample approximations for the distributions of the SR and the PSR estimators without knowing the values of unknown parameters.

Also, as shown in Appendix A, the asymptotic distribution of the SR estimator is given as follows:

$$\sqrt{n}(\widehat{\mu}_{\mathrm{SR}} - \mu) \xrightarrow{d} \begin{cases} \left(1 - \dfrac{k-2}{Z'Z}\right) Z & \text{if } \mu = 0, \\[2ex] Z & \text{if } \mu \neq 0, \end{cases} \tag{3}$$

where $Z \sim \mathrm{N}(0, \mathrm{I}_k)$. Similarly, the asymptotic distribution of the PSR estimator is given by

$$\sqrt{n}(\widehat{\mu}_{\mathrm{PSR}} - \mu) \xrightarrow{d} \begin{cases} \max\left[0,\ 1 - \dfrac{k-2}{Z'Z}\right] Z & \text{if } \mu = 0, \\ Z & \text{if } \mu \neq 0. \end{cases} \tag{4}$$

From the above results, we can see that the distributions of the SR and the PSR estimators are not continuous at $\mu = 0$.

3 Bootstrap methods

When the distribution of a statistic is complex or unknown, the bootstrap method proposed by Efron (1979) is often useful. In this section, we analytically investigate the validity of the bootstrap methods for the SR and the PSR estimators. Thus, the aim of this section is to approximate the asymptotic distributions of the SR and the PSR estimators given in (3) and (4) by bootstrap methods. Hereafter, we denote the bootstrap sample of size m (i.e., the random sample of size m obtained by sampling with replacement from the original sample $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$) by $\mathcal{X}_m^* = \{X_1^*, X_2^*, \ldots, X_m^*\}$.

3.1 Conventional bootstrap

In the conventional bootstrap, the distribution of a statistic is approximated by the distribution of the statistic obtained by replacing the sample in the original statistic with the bootstrap sample of size n. Thus, the distribution of the SR estimator is approximated by the distribution of

$$\widehat{\mu}_{\mathrm{SR}}^* = \left(1 - \frac{k-2}{n\overline{X}^{*\prime}\overline{X}^*}\right)\overline{X}^*, \tag{5}$$

where $\overline{X}^* = \frac{1}{n}\sum_{i=1}^n X_i^*$ is the sample mean of the bootstrap sample of size n. Similarly, the distribution of the PSR estimator is approximated by the distribution of

$$\widehat{\mu}_{\mathrm{PSR}}^* = \max\left[0,\ 1 - \frac{k-2}{n\overline{X}^{*\prime}\overline{X}^*}\right]\overline{X}^*. \tag{6}$$

As shown in Appendix B, the asymptotic distributions of $\widehat{\mu}_{\mathrm{SR}}^*$ and $\widehat{\mu}_{\mathrm{PSR}}^*$ are given by

$$\sqrt{n}(\widehat{\mu}_{\mathrm{SR}}^* - \overline{X}) \overset{d}{\approx} \begin{cases} \left(1 - \dfrac{k-2}{(Z^*+Z)'(Z^*+Z)}\right) Z^* - \dfrac{k-2}{(Z^*+Z)'(Z^*+Z)} Z & \text{if } \mu = 0, \\ Z^* & \text{if } \mu \neq 0, \end{cases} \tag{7}$$

and

$$\sqrt{n}(\widehat{\mu}_{\mathrm{PSR}}^* - \overline{X}) \overset{d}{\approx} \begin{cases} \max\left[0,\ 1 - \dfrac{k-2}{(Z^*+Z)'(Z^*+Z)}\right] Z^* & \\ \quad + \max\left[-1,\ -\dfrac{k-2}{(Z^*+Z)'(Z^*+Z)}\right] Z & \text{if } \mu = 0, \\ Z^* & \text{if } \mu \neq 0, \end{cases} \tag{8}$$

where $Z^* \sim \mathrm{N}(0, \mathrm{I}_k)$, $Z \sim \mathrm{N}(0, \mathrm{I}_k)$, and $\overset{d}{\approx}$ denotes that the asymptotic distribution of the statistic on the left hand side is given by the distribution of the random variable in the right hand side, given a realization of Z. Note that Z is considered as a realization from $\mathrm{N}(0, \mathrm{I}_k)$ since the bootstrap is executed given $\mathcal{X}$. For example, "$W \overset{d}{\approx} Z^* + Z$" means that "the asymptotic distribution of W is the distribution of $Z^* + Z$ given a realization of Z", i.e., "$\mathrm{N}(Z, I_k)$ given Z". Therefore, the original sample $\mathcal{X}$ affects the asymptotic bootstrap distributions of SR and PSR estimators through a realization of Z when $\mu = 0$. Thus, as discussed by Beran (1997), the bootstrap distributions of the SR and PSR estimators are random probability measures.

Comparing equations (3), (4), (7) and (8), we can see that the conventional bootstrap is consistent for the SR estimator if $\mu \neq 0$ and it is not consistent if $\mu = 0$. Thus, it is expected that the conventional bootstrap does not work around $\mu = 0$. Similarly, the convectional bootstrap for the PSR estimator is consistent if $\mu \neq 0$ and not consistent if $\mu = 0$.

3.2 Centered bootstrap

In the previous subsection, we see that the conventional bootstrap for the SR and the PSR estimators is not consistent if $\mu = 0$. Thus, in this subsection, we consider to approximate the distributions of the SR and PSR estimators by those of the centered variants:

$$\tilde{\mu}^*_{\mathrm{SR}} = \left(1 - \frac{k-2}{n(\overline{X}^* - \overline{X})'(\overline{X}^* - \overline{X})}\right)(\overline{X}^* - \overline{X}) + \overline{X}, \tag{9}$$

$$\tilde{\mu}^*_{\mathrm{PSR}} = \max\left[0,\ 1 - \frac{k-2}{n(\overline{X}^* - \overline{X})'(\overline{X}^* - \overline{X})}\right](\overline{X}^* - \overline{X}) + \overline{X}. \tag{10}$$

We call this method the centered bootstrap. As shown in Appendix B, the asymptotic distributions of the centered bootstrap estimators are given by

$$\sqrt{n}(\tilde{\mu}^*_{\mathrm{SR}} - \overline{X}) \xrightarrow{d} \left(1 - \frac{k-2}{Z^{*\prime}Z^*}\right)Z^*, \tag{11}$$

$$\sqrt{n}(\tilde{\mu}^*_{\mathrm{PSR}} - \overline{X}) \xrightarrow{d} \max\left[0,\ 1 - \frac{k-2}{Z^{*\prime}Z^*}\right]Z^*. \tag{12}$$

Thus, comparing equations (3), (4), (11) and (12), the centered bootstrap for the SR and the PSR estimators is consistent for $\mu = 0$ and it is not consistent for $\mu \neq 0$.

3.3 m out of n bootstrap

So far, we have been considering to use the bootstrap sample of size n. However, in this section, we consider to utilize the bootstrap sample of size m where m satisfies $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Following Bickel et al. (1997), Bickel and Sakov (2008) and Wei et al.(2016), we call this method the m out of n bootstrap. In the m out of n bootstrap, the distributions of the SR and the PSR estimators are approximated by those of

$$\hat{\mu}^\dagger_{\mathrm{SR}} = \left(1 - \frac{k-2}{m\overline{X}^{\dagger\prime}\overline{X}^\dagger}\right)\overline{X}^\dagger, \tag{13}$$

and

$$\widehat{\mu}^{\dagger}_{\mathrm{PSR}} = \max\left[0,\ 1 - \frac{k-2}{m\overline{X}^{\dagger\prime}\overline{X}^{\dagger}}\right]\overline{X}^{\dagger}, \tag{14}$$

where $\overline{X}^{\dagger} = \frac{1}{m}\sum_{i=1}^{m} X_i^*$ is the sample mean of the bootstrap sample of size m. As shown in Appendix B, the asymptotic distributions of $\widehat{\mu}^{\dagger}_{\mathrm{SR}}$ and $\widehat{\mu}^{\dagger}_{\mathrm{PSR}}$ are given by

$$\sqrt{m}(\widehat{\mu}^{\dagger}_{\mathrm{SR}} - \overline{X}) \xrightarrow{d} \begin{cases} \left(1 - \dfrac{k-2}{Z^{*\prime}Z^*}\right)Z^* & \text{if } \mu = 0, \\ Z^* & \text{if } \mu \neq 0, \end{cases} \tag{15}$$

and

$$\sqrt{m}(\widehat{\mu}^{\dagger}_{\mathrm{PSR}} - \overline{X}) \xrightarrow{d} \begin{cases} \max\left[0,\ 1 - \dfrac{k-2}{Z^{*\prime}Z^*}\right]Z^* & \text{if } \mu = 0, \\ Z^* & \text{if } \mu \neq 0, \end{cases} \tag{16}$$

where $Z^* \sim \mathrm{N}(0, \mathrm{I}_k)$. Comparing equations (3), (4), (15) and (16), we can see that the m out of n bootstrap is consistent for both $\mu = 0$ and $\mu \neq 0$. To apply the m out of n bootstrap, we have to decide the value of m. Bickel and Sakov (2008) and Wei et al. (2016) proposed the ways to choose m and investigated their performances. They showed that the m chosen by their ways converges to n (i.e., the conventional bootstrap) when the conventional bootstrap is consistent, and it converges to the optimal m in the sense that the distance between the approximated distribution and the true distribution is minimized.

3.4 Bootstrap based on a pre-test

As shown in the previous subsection, the m out of n bootstrap enables us to approximate the distributions of the SR and PSR estimators consistently. However, the distributions of SR and PSR estimators depends on the sample size n. Since the different sample size m from n is used in m out of n bootstrap, the approximate distributions may be far from exact distributions. Thus, the same sample size n as the original sample size may be preferred in bootstrap. Thus, according to the idea introduced in the parametric bootstrap by Beran (1997), we consider the following nonparametric bootstrap which incorporates a pre-test.

As we showed above the centered bootstrap is consistent if $\mu = 0$, and the conventional bootstrap is consistent if $\mu \neq 0$. Thus, we consider to use the centered bootstrap if μ can be considered as 0, and use the conventional bootstrap otherwise. Using the fact that $n\overline{X}'\overline{X} = O_p(1)$ if $\mu = 0$ and $n\overline{X}'\overline{X} = O_p(n)$ if $\mu \neq 0$, we consider to approximate the distributions of the SR and PSR estimators by those of

$$I(|\overline{X}| < n^{\gamma})\tilde{\mu}^*_{\mathrm{SR}} + \{1 - I(|\overline{X}| < n^{\gamma})\}\widehat{\mu}^*_{\mathrm{SR}},$$

$$I(|\overline{X}| < n^{\gamma})\tilde{\mu}^*_{\mathrm{PSR}} + \{1 - I(|\overline{X}| < n^{\gamma})\}\widehat{\mu}^*_{\mathrm{PSR}},$$

where $I(A)$ is an indicator function such that $I(A) = 1$ if an event A occurs and $I(A) = 0$ otherwise. We call this method "pre-test bootstrap" hereafter. When $-1/2 < \gamma < 0$, $I(|\overline{X}| <$

$n^{\gamma}) \xrightarrow{p} 1$ for $\mu = 0$ and $I(|\overline{X}| < n^{\gamma}) \xrightarrow{p} 0$ for $\mu \neq 0$. Thus, it is easy to show that the pre-test bootstrap is consistent for both $\mu = 0$ and $\mu \neq 0$.

Though we have to decide the value of m before applying the m out of n bootstrap, we do not have to make such a decision in the pre-test bootstrap. We have to choose γ, the order of the critical value of the pre-test instead.

As discussed above, the conventional bootstrap is consistent only for $\mu \neq 0$ while the centered bootstrap is consistent only for $\mu = 0$. On the other hand, the m out of n bootstrap and the pre-test bootstrap are consistent for both $\mu = 0$ and $\mu \neq 0$. To see the performances of these methods, we execute Monte Carlo simulations in the next section.

4 Simulation results

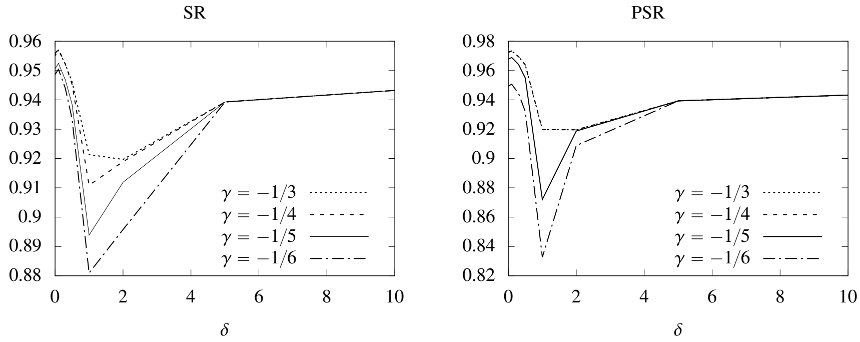
To investigate the performances of the methods proposed in the previous section, we execute some Monte Carlo simulations in this section. In the simulations, we take the dimension of the normal distribution $k = 3, 4, 5, 7, 10$. Since the conventional bootstrap is not consistent for $\mu = 0$ and the centered bootstrap is not consistent for $\mu \neq 0$, we are interested in the performances of these methods around $\mu = 0$. Thus, we considered the case where $\mu = \delta\iota/\sqrt{n}$, ι is the $k \times 1$ vector whose elements are 1s (i.e., we consider the case such that all the elements of μ are equal), and $\delta = 0, 0.1, 0.3, 0.5, 1, 2, 5, 10, 15, 20, 25$. Based on the above setup of the mean vector μ, random vectors X_i from $\mathrm{N}(\mu, \mathrm{I_k})$, where the sample size $n = 20, 50, 100, 200, 500$ and 1000 are generated. Note that every method is inconsistent when $\mu = \delta\iota/\sqrt{n}$ and $\delta \neq 0$, since there is no way to obtain a consistent estimator for δ.

In order to investigate the accuracy of the method, we constructed 90%, 95% and 99% confidence intervals for μ_i (i.e., the ith element of μ) by the approximated distributions of the SR and PSR estimators, and calculated the ratio such that the confidence intervals contain the true mean $\mu_i = \delta/\sqrt{n}$, based on 100000 Monte Carlo experiments. In constructing the bootstrap approximations, we executed 500 bootstrap iterations for each method.

First, to investigate the effect of the order of the pre-test critical value γ, we executed simulations for various values of $\gamma \in (-1/2, 0)$. Figure 1 shows the coverages of the 95% confidence intervals obtained by the pre-test bootstrap for $\gamma = -1/3, -1/4, -1/5$ and $-1/6$ when $k = 7$ and $n = 100$. We can see that the effect of γ is very small when δ is far from 0. When δ is around 1, the coverages for $\gamma = -1/5$ and $-1/6$ are far from nominal coverage (i.e., 0.95). Also, as a whole the performance of the pre-test bootstrap confidence intervals for $\gamma = -1/3$ and $-1/4$ are almost comparable. Though we do not show the results for the other cases, we obtained similar results for all the cases. Thus, following Beran (1997), we use $\gamma = -1/4$ in the following simulations.

Table 1 shows the results for the case of $k = 4$ and $n = 50$. In the tables, 'Centered BS', 'Conv BS' and 'Pre-test' indicate the centered bootstrap, conventional bootstrap and the pre-test bootstrap. '$m = \lfloor n^{\alpha} \rfloor$' shows the results for the m out of n bootstrap with sample size m. Also, for a purpose of comparison, we calculated the ratio such that the confidence intervals based on normal approximation of the distributions of the SR and the PSR estimators contain the true mean, and 'Normal Approx' shows the results obtained by this normal approximation. In Tables 1 – 4, the values closest to the nominal coverage are highlighted. We can see from Table 1 that the centered bootstrap, m out of n bootstrap and the pre-test bootstrap work well for $\mu = 0$. Also, the normal approximation does not work for $\mu = 0$. These results coincide with the theoretical results. However, it seems that the conventional bootstrap yields rather accurate confidence intervals for $\mu = 0$ except for the 90% confi-

Fig. 1 Coverage of the 95% confidence intervals based on the pre-test bootstrap



dence interval of the PSR estimator, though the conventional bootstrap is inconsistent as shown in the previous section. We will give discussion about this point later.

When μ is close to 0 but not equal to 0 (i.e., $\mu_i = 1/\sqrt{n}$), every method is not so accurate. This is because the asymptotic distributions of the estimators are not continuous at $\mu = 0$. In this case, conventional bootstrap seems to be most reliable among the methods considered here. The approximation of the m out of n bootstrap gets better as m increases. Also, the approximation of the pre-test bootstrap is almost comparable to the approximation of the m out of n bootstrap with $m = \lfloor n^{2/3} \rfloor$[1].)

When μ is far from 0 (i.e., $\mu_i = 25/\sqrt{n}$), the conventional bootstrap, pre-test bootstrap and the normal approximation work well. In this case, the results of the pre-test bootstrap are the same as those of the conventional bootstrap. However, the centered bootstrap is not valid. These results coincide with the theoretical results. Though the theoretical result shows that the m out of n bootstrap is valid for $\mu \neq 0$, it is not so accurate as the conventional bootstrap when small values of m are used. We can see that the results of the m out of n bootstrap approach to those of the conventional bootstrap as m increases. This is quite natural that the m out of n bootstrap reduces to the conventional bootstrap when $m = n$.

Table 2 shows the results for $k = 4$ and $n = 100$. From Table 2, we can see the similar results to Table 1. Comparing Tables 1 and 2, we can see that the effect of the sample size n is not so large. Also, in most cases, it seems that the effect of the choice of m is not very large though larger m seems to be preferred. When $m = \lfloor n^{1/3} \rfloor$ is used, m is very small (i.e., $m = 3$ for $n = 50$ and $m = 4$ for $n = 100$). Nevertheless, the results for $m = \lfloor n^{1/3} \rfloor$ are not so inaccurate compared with the results for the other choice of m for $\mu_i = 0$ and $\mu_i = 25/\sqrt{n}$.

Tables 3 and 4 show the results for $k = 7$, $n = 50$ and 100. We can see that the centered bootstrap is valid for $\mu = 0$. The approximation of m out of n bootstrap is not so accurate when m is small (i.e., $m = \lfloor n^{1/3} \rfloor$). In contrast to the case of $k = 4$, the conventional bootstrap does not work for $\mu = 0$, which is consistent with the theoretical results. Noting that $\mathrm{E}[1/n\overline{X}'\overline{X}] = 1/k$, the effect of shrinkage factor $\left(1 - \frac{k-2}{n\overline{X}'\overline{X}}\right)$ will be small for small k. Thus, even when $\mu = 0$, the conventional bootstrap may seem to work for small values of k. Also, because of the inaccuracy of the conventional bootstrap, the pre-test bootstrap for the PSR estimator does not work very well for $\mu = 0$ (e.g., the coverage of 90% confidence interval is 0.936 for $n = 50$). On the other hand, the pre-test bootstrap works very well for the SR estimator for $\mu = 0$.

[1] Wei et al (2016) proposed to use $m \propto n^{2/3}$ under the assumption $\mu = o(n^{-2/3})$.

Table 1 Results for $k = 4$ and $n = 50$.

		SR			PSR		
μ_i	method	99%	95%	90%	99%	95%	90%
0	Centered BS	0.987	0.945	0.894	0.984	0.941	0.888
	$m = \lfloor n^{1/3} \rfloor$	0.982	0.936	0.881	0.983	0.942	0.891
	$m = \lfloor n^{1/2} \rfloor$	0.982	0.938	0.884	0.986	0.947	0.898
	$m = \lfloor n^{2/3} \rfloor$	0.984	0.941	0.888	0.988	0.951	0.905
	$m = \lfloor n^{3/4} \rfloor$	0.985	0.943	0.892	0.989	0.954	0.910
	$m = \lfloor n^{4/5} \rfloor$	0.985	0.945	0.895	0.989	0.956	0.913
	Conv BS	0.987	0.953	0.910	0.991	0.963	0.926
	Pre-test	0.986	0.950	0.901	0.990	0.954	0.903
	Norm Approx	0.995	0.982	0.963	0.998	0.988	0.973
$1/\sqrt{n}$	Centered BS	0.985	0.921	0.818	0.975	0.894	0.755
	$m = \lfloor n^{1/3} \rfloor$	0.977	0.908	0.808	0.975	0.899	0.769
	$m = \lfloor n^{1/2} \rfloor$	0.979	0.913	0.817	0.979	0.911	0.789
	$m = \lfloor n^{2/3} \rfloor$	0.981	0.918	0.828	0.982	0.920	0.809
	$m = \lfloor n^{3/4} \rfloor$	0.982	0.923	0.835	0.983	0.926	0.823
	$m = \lfloor n^{4/5} \rfloor$	0.983	0.925	0.840	0.984	0.929	0.832
	Conv BS	0.986	0.938	0.867	0.987	0.944	0.870
	Pre-test	0.985	0.934	0.845	0.986	0.928	0.811
	Norm Approx	0.996	0.977	0.949	0.997	0.981	0.956
$25/\sqrt{n}$	Centered BS	0.969	0.859	0.747	0.943	0.812	0.686
	$m = \lfloor n^{1/3} \rfloor$	0.980	0.936	0.884	0.980	0.936	0.884
	$m = \lfloor n^{1/2} \rfloor$	0.982	0.938	0.887	0.982	0.938	0.887
	$m = \lfloor n^{2/3} \rfloor$	0.983	0.939	0.888	0.983	0.939	0.888
	$m = \lfloor n^{3/4} \rfloor$	0.983	0.940	0.888	0.983	0.940	0.888
	$m = \lfloor n^{4/5} \rfloor$	0.983	0.940	0.888	0.983	0.940	0.888
	Conv BS	0.984	0.940	0.888	0.984	0.940	0.888
	Pre-test	0.984	0.940	0.888	0.984	0.940	0.888
	Norm Approx	0.990	0.950	0.900	0.990	0.950	0.900

As for the sample sizes m in m out of n bootstrap, $\lfloor n^{1/3} \rfloor = 3$, $\lfloor n^{1/2} \rfloor = 7$, $\lfloor n^{2/3} \rfloor = 13$, $\lfloor n^{3/4} \rfloor = 18$, $\lfloor n^{4/5} \rfloor = 22$ are used.

For the other aspects, we can see the similar results to the case of $k = 4$. The centered bootstrap works very well and the normal approximation is not valid for $\mu = 0$. No method considered in this paper is very accurate for small μ (i.e., $\mu_i = 1/\sqrt{n}$). When μ is far from 0, the conventional bootstrap works well and the pre-test bootstrap has the same coverage as the conventional bootstrap, however, the centered bootstrap is not valid. These results coincide with the theoretical results obtained in the previous section. Also, it seems that the effects of the sample size n and the choice of m in the m out of n bootstrap are not so large. We can also see that the results of the m out of n bootstrap approaches to the ones obtained by the conventional bootstrap as m increases.

As a whole, the centered bootstrap seems to be the best choice when $\mu = 0$ though the m out of n bootstrap is also consistent when $\mu = 0$. Also, the conventional bootstrap yields the better performance than the m out of n bootstrap for $\mu \neq 0$ though both the m out of n bootstrap and the conventional bootstrap are consistent when $\mu \neq 0$. When μ is not 0 but close to 0, each method is not so accurate. However, the conventional bootstrap is more reliable than the other methods. The performances of the pre-test bootstrap and the m out of n bootstrap with large values of m are not so bad compared with the performances of the

Table 2 Results for $k = 4$ and $n = 100$.

		SR			PSR		
μ_i	method	99%	95%	90%	99%	95%	90%
0	Centered BS	0.988	0.947	0.895	0.987	0.945	0.893
	$m = \lfloor n^{1/3} \rfloor$	0.985	0.941	0.888	0.987	0.947	0.896
	$m = \lfloor n^{1/2} \rfloor$	0.985	0.942	0.888	0.988	0.951	0.901
	$m = \lfloor n^{2/3} \rfloor$	0.986	0.945	0.892	0.990	0.955	0.908
	$m = \lfloor n^{3/4} \rfloor$	0.986	0.947	0.896	0.991	0.958	0.913
	$m = \lfloor n^{4/5} \rfloor$	0.987	0.948	0.898	0.991	0.959	0.916
	Conv BS	0.988	0.956	0.914	0.993	0.966	0.930
	Pre-test	0.988	0.949	0.898	0.989	0.949	0.898
	Norm Approx	0.995	0.981	0.964	0.998	0.988	0.973
$1/\sqrt{n}$	Centered BS	0.987	0.922	0.823	0.979	0.904	0.767
	$m = \lfloor n^{1/3} \rfloor$	0.982	0.915	0.814	0.980	0.909	0.778
	$m = \lfloor n^{1/2} \rfloor$	0.983	0.918	0.820	0.983	0.917	0.794
	$m = \lfloor n^{2/3} \rfloor$	0.984	0.923	0.831	0.985	0.926	0.816
	$m = \lfloor n^{3/4} \rfloor$	0.985	0.927	0.840	0.986	0.932	0.831
	$m = \lfloor n^{4/5} \rfloor$	0.986	0.930	0.845	0.987	0.935	0.840
	Conv BS	0.988	0.943	0.875	0.990	0.949	0.882
	Pre-test	0.988	0.935	0.845	0.987	0.929	0.806
	Norm Approx	0.996	0.977	0.949	0.997	0.981	0.956
$25/\sqrt{n}$	Centered BS	0.969	0.860	0.750	0.948	0.820	0.696
	$m = \lfloor n^{1/3} \rfloor$	0.983	0.939	0.888	0.983	0.939	0.888
	$m = \lfloor n^{1/2} \rfloor$	0.985	0.942	0.891	0.985	0.942	0.891
	$m = \lfloor n^{2/3} \rfloor$	0.985	0.942	0.892	0.985	0.942	0.892
	$m = \lfloor n^{3/4} \rfloor$	0.985	0.943	0.892	0.985	0.943	0.892
	$m = \lfloor n^{4/5} \rfloor$	0.986	0.943	0.892	0.986	0.943	0.892
	Conv BS	0.986	0.944	0.893	0.986	0.944	0.893
	Pre-test	0.986	0.944	0.893	0.986	0.944	0.893
	Norm Approx	0.990	0.949	0.900	0.990	0.949	0.900

As for the sample sizes m in m out of n bootstrap, $\lfloor n^{1/3} \rfloor = 4$, $\lfloor n^{1/2} \rfloor = 10$, $\lfloor n^{2/3} \rfloor = 21$, $\lfloor n^{3/4} \rfloor = 31$, $\lfloor n^{4/5} \rfloor = 39$ are used.

centered bootstrap, m out of n bootstrap with small value of m and the normal approximation when μ is close to 0. Considering that the conventional bootstrap is not consistent when $\mu = 0$ and that the m out of n bootstrap approaches to the conventional bootstrap as m increases to n, the pre-test bootstrap may be a reliable method.

Next, we investigate the lengths of the confidence intervals obtained by the proposed methods. We calculated the "relative length" defined as the average length of the confidence interval obtained by the proposed method divided by the length of the confidence interval based on the conventional MLE. Thus, a confidence interval has shorter average length than the MLE-based confidence interval if the value of relative length is smaller than 1. Figure 2 shows the relative lengths of the 90% confidence intervals based on the m out of n bootstrap and the pre-test bootstrap for the SR and PSR estimators. Because only the m out of n bootstrap and the pre-test bootstrap are consistent for both $\mu = 0$ and $\mu \neq 0$, we show the results only for these methods. We can see that the confidence intervals by the bootstrap methods for the SR and the PSR estimators are much shorter than the MLE-based confidence interval when δ is close to 0. Considering that the coverage of the confidence interval by the m out of n bootstrap with $m = \lfloor n^{1/3} \rfloor$ is much smaller than the nominal level, it is

Table 3 Results for $k = 7$ and $n = 50$.

		SR			PSR		
μ_i	method	99%	95%	90%	99%	95%	90%
0	Centered BS	0.989	0.950	0.902	0.985	0.943	0.890
	$m = \lfloor n^{1/3} \rfloor$	0.981	0.930	0.872	0.983	0.937	0.880
	$m = \lfloor n^{1/2} \rfloor$	0.983	0.934	0.874	0.986	0.944	0.888
	$m = \lfloor n^{2/3} \rfloor$	0.985	0.941	0.883	0.988	0.950	0.899
	$m = \lfloor n^{3/4} \rfloor$	0.987	0.945	0.890	0.990	0.954	0.907
	$m = \lfloor n^{4/5} \rfloor$	0.987	0.948	0.897	0.990	0.957	0.913
	Conv BS	0.991	0.962	0.923	0.994	0.969	0.935
	Pre-test	0.989	0.953	0.907	0.994	0.969	0.936
	Norm Approx	0.998	0.993	0.985	1.000	0.997	0.991
$1/\sqrt{n}$	Centered BS	0.984	0.871	0.693	0.960	0.768	0.554
	$m = \lfloor n^{1/3} \rfloor$	0.965	0.836	0.676	0.955	0.783	0.597
	$m = \lfloor n^{1/2} \rfloor$	0.970	0.851	0.702	0.967	0.821	0.646
	$m = \lfloor n^{2/3} \rfloor$	0.974	0.869	0.730	0.974	0.855	0.697
	$m = \lfloor n^{3/4} \rfloor$	0.977	0.881	0.751	0.977	0.874	0.729
	$m = \lfloor n^{4/5} \rfloor$	0.979	0.889	0.765	0.979	0.886	0.750
	Conv BS	0.986	0.926	0.834	0.987	0.928	0.836
	Pre-test	0.986	0.914	0.791	0.986	0.887	0.759
	Norm Approx	0.999	0.989	0.970	0.999	0.990	0.973
$25/\sqrt{n}$	Centered BS	0.932	0.761	0.625	0.873	0.671	0.519
	$m = \lfloor n^{1/3} \rfloor$	0.979	0.933	0.880	0.979	0.933	0.880
	$m = \lfloor n^{1/2} \rfloor$	0.982	0.938	0.884	0.982	0.938	0.884
	$m = \lfloor n^{2/3} \rfloor$	0.982	0.939	0.886	0.982	0.939	0.886
	$m = \lfloor n^{3/4} \rfloor$	0.983	0.939	0.887	0.983	0.939	0.887
	$m = \lfloor n^{4/5} \rfloor$	0.983	0.939	0.887	0.983	0.939	0.887
	Conv BS	0.984	0.940	0.887	0.984	0.940	0.887
	Pre-test	0.984	0.940	0.887	0.984	0.940	0.887
	Norm Approx	0.990	0.950	0.900	0.990	0.950	0.900

As for the sample sizes m in m out of n bootstrap, $\lfloor n^{1/3} \rfloor = 3$, $\lfloor n^{1/2} \rfloor = 7$, $\lfloor n^{2/3} \rfloor = 13$, $\lfloor n^{3/4} \rfloor = 18$, $\lfloor n^{4/5} \rfloor = 22$ are used.

natural that the confidence interval with $m = \lfloor n^{1/3} \rfloor$ is shortest. As a whole, we can see that it is possible to construct shorter confidence intervals than the conventional MLE-based confidence interval by using the SR and the PSR estimators.

Also, in order to examine the performances of the proposed methods under more general situations, we executed some additional Monte Carlo experiments for more general cases. Figure 3 shows the results for $k = 7$ and $n = 100$. Since these results are typical, we do not show the results for other values of k and n. Figure 3 shows the empirical coverages of 90% confidence intervals for the first element of the mean vector based on the SR and the PSR estimators using the m out of n bootstrap with $m = \lfloor n^{2/3} \rfloor$ (denoted as m-boot) and the pre-test bootstrap (denoted as pre-test). Note that they are consistent both for $\mu = 0$ and $\mu \neq 0$. Since the consistency of the other methods depends on whether $\mu = 0$ or $\mu \neq 0$, the coverages of them are not depicted. The setups of the experiments are as follows:

Benchmark: X_i's are generated in the same way as in the one used in Table 4. This result is shown in the purpose of comparison with the other cases. In this case the noncentrality parameter is given by $\mu'\mu = k\delta^2/n$.

Table 4 Results for $k = 7$ and $n = 100$.

		SR			PSR		
μ_i	method	99%	95%	90%	99%	95%	90%
0	Centered BS	0.989	0.950	0.900	0.987	0.945	0.894
	$m = \lfloor n^{1/3} \rfloor$	0.984	0.937	0.878	0.986	0.942	0.886
	$m = \lfloor n^{1/2} \rfloor$	0.985	0.937	0.877	0.988	0.946	0.891
	$m = \lfloor n^{2/3} \rfloor$	0.987	0.942	0.884	0.990	0.953	0.901
	$m = \lfloor n^{3/4} \rfloor$	0.988	0.947	0.892	0.991	0.957	0.908
	$m = \lfloor n^{4/5} \rfloor$	0.989	0.950	0.898	0.992	0.959	0.914
	Conv BS	0.992	0.966	0.927	0.994	0.973	0.939
	Pre-test	0.990	0.957	0.912	0.994	0.968	0.925
	Norm Approx	0.999	0.993	0.985	1.000	0.997	0.991
$1/\sqrt{n}$	Centered BS	0.982	0.869	0.688	0.963	0.781	0.562
	$m = \lfloor n^{1/3} \rfloor$	0.971	0.847	0.681	0.964	0.793	0.591
	$m = \lfloor n^{1/2} \rfloor$	0.972	0.856	0.697	0.969	0.823	0.632
	$m = \lfloor n^{2/3} \rfloor$	0.976	0.870	0.725	0.976	0.856	0.686
	$m = \lfloor n^{3/4} \rfloor$	0.979	0.883	0.748	0.979	0.878	0.723
	$m = \lfloor n^{4/5} \rfloor$	0.981	0.892	0.765	0.981	0.890	0.747
	Conv BS	0.988	0.931	0.842	0.989	0.934	0.844
	Pre-test	0.987	0.911	0.768	0.987	0.871	0.712
	Norm Approx	0.999	0.989	0.969	0.999	0.990	0.972
$25/\sqrt{n}$	Centered BS	0.928	0.758	0.620	0.879	0.678	0.526
	$m = \lfloor n^{1/3} \rfloor$	0.982	0.935	0.883	0.982	0.935	0.883
	$m = \lfloor n^{1/2} \rfloor$	0.984	0.940	0.888	0.984	0.940	0.888
	$m = \lfloor n^{2/3} \rfloor$	0.985	0.942	0.891	0.985	0.942	0.891
	$m = \lfloor n^{3/4} \rfloor$	0.986	0.943	0.891	0.986	0.943	0.891
	$m = \lfloor n^{4/5} \rfloor$	0.985	0.943	0.892	0.985	0.943	0.892
	Conv BS	0.987	0.944	0.893	0.987	0.944	0.893
	Pre-test	0.987	0.944	0.893	0.987	0.944	0.893
	Norm Approx	0.990	0.950	0.900	0.990	0.950	0.900

As for the sample sizes m in m out of n bootstrap, $\lfloor n^{1/3} \rfloor = 4$, $\lfloor n^{1/2} \rfloor = 10$, $\lfloor n^{2/3} \rfloor = 21$, $\lfloor n^{3/4} \rfloor = 31$, $\lfloor n^{4/5} \rfloor = 39$ are used.

Sparse 1: X_i's are generated as the normal random vector whose first element of the mean vector μ is $\sqrt{k}\delta/\sqrt{n}$ and the other elements are zeros. In this case, we have the same noncentrality parameter $\mu'\mu = k\delta^2/n$ as in the Benchmark case.

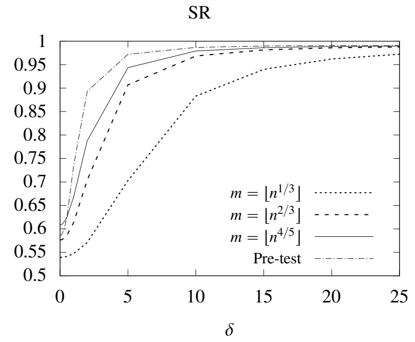
Sparse 2: The kth element of the mean vector is $\sqrt{k}\delta/\sqrt{n}$ and the other elements are zeros. Note that the first element of the true mean for which we are constructing the confidence intervals is fixed at zero. Similar to the former two setups, the noncentrality parameter is $\mu'\mu = k\delta^2/n$.

General Covariance: In the previous cases, we assume that the covariance matrix of X_i is given by I_k. Here, we will examine the effect of the departure of the covariance matrix from the identity matrix. When the covariance matrix of X_i is given by unknown Σ, the SR and PSR estimators are defined by

$$\hat{\mu}_{\mathrm{SR}} = \left(1 - \frac{k-2}{n\overline{X}'\hat{\Sigma}^{-1}\overline{X}}\right)\overline{X}, \quad \hat{\mu}_{\mathrm{PSR}} = \max\left[0,\ 1 - \frac{k-2}{n\overline{X}'\hat{\Sigma}^{-1}\overline{X}}\right]\overline{X}, \tag{17}$$

where $\hat{\Sigma} = n^{-1}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})'$ is a consistent estimator for Σ. In order to investigate the performances for these estimators, generating a $k \times k$ whose elements

Fig. 2 Lengths of the 90% confidence intervals based on the SR and PSR estimators



are random draws from $U(0, 1)$ and dividing its elements by the sum of each row elements, we obtained a matrix A. Fixing the matrix A, we generate a random vector $X_i = A(Z_i + \mu)$, where $Z_i \sim \mathrm{N}(0, I_k)$ and $\mu = \delta\iota/\sqrt{n}$. Thus, $X_i \sim \mathrm{N}(A\mu, AA')$. Because of the structure of A, we have $A\mu = \mu$. Therefore, the noncentrality parameter is given by $(A\mu)'\Sigma^{-1}A\mu = \mu' A'(AA')^{-1}A\mu = \mu'\mu = k\delta^2/n$, which is the same value as in the previous three setups. Since the noncentrality parameters are the same for all setups, we can compare the performances of the methods based on the values of δ. Showing that the modifications of the bootstrap variants of the estimators like

$$\widetilde{\mu}^*_{\mathrm{SR}} = \left(1 - \frac{k-2}{n(\overline{X}^* - \overline{X})'\widehat{\Sigma}^{*-1}(\overline{X}^* - \overline{X})}\right)(\overline{X}^* - \overline{X}) + \overline{X}, \tag{18}$$

$$\widehat{\mu}^\dagger_{\mathrm{SR}} = \left(1 - \frac{k-2}{m\overline{X}^{\dagger\prime}\widehat{\Sigma}^{*-1}\overline{X}^\dagger}\right)\overline{X}^\dagger, \tag{19}$$

yield the similar theoretical results is a straightforward extension of the proofs given in Appendices. Since any consistent estimator $\widehat{\Sigma}^*$ for Σ can be used in the above equations, we used a analogue of $\widehat{\Sigma}$ based on the bootstrap sample as $\widehat{\Sigma}^*$.

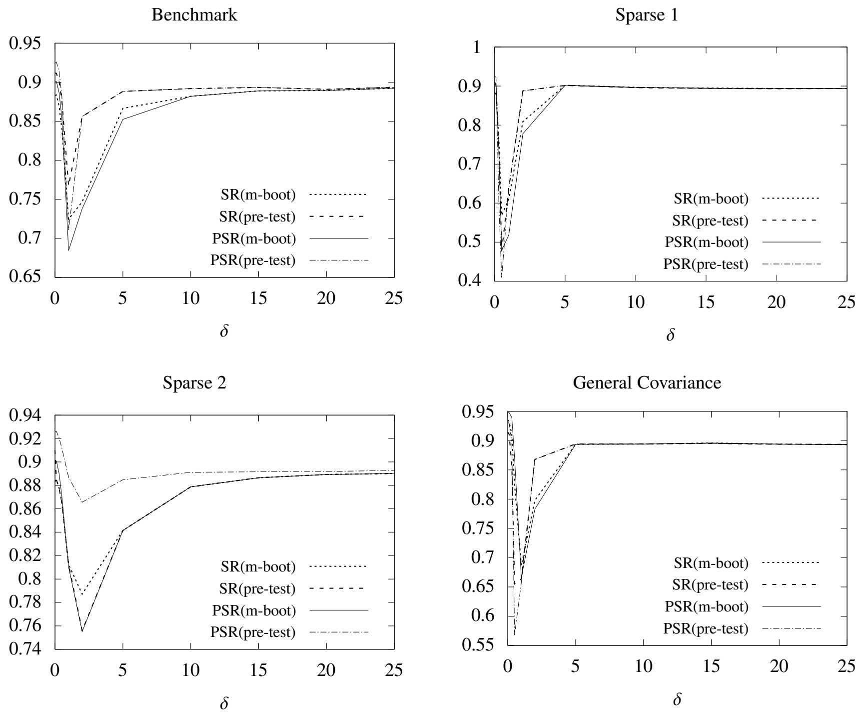
As we can see Figure 3, all the methods are accurate at $\delta = 0$ and δ is far from 0. However, each method is not so accurate when δ is close to 0. We can see that the coverage of the pre-test bootstrap for the PSR estimator can be very when δ is close to 0 but not 0. However, the pre-test bootstrap is more accurate than the m out of n bootstrap at $\delta = 0$. Though empirical coverages in ‘Sparse 1’ case can be very far from the nominal level compared to the case of ‘Sparse 2’, overall behavior is almost similar. ‘General Covariance’ in Figure 3 shows the similar results to the ‘Benchmark’. This implies that the proposed methods can be extended to the cases of general covariance matrices.

Thus, though there are several little differences depending on setups and parameter values, we can see that the proposed methods are applicable under rather general situations.

5 Concluding remarks

In this paper, we considered to apply the bootstrap methods to the SR and PSR estimators for the mean of a normal distribution. We analytically showed that the conventional bootstrap

Fig. 3 Coverages of the 90% confidence intervals



method is not consistent when the mean of the distribution μ is 0 though it is consistent when μ is not 0. Thus, we proposed the centered bootstrap which is consistent when $\mu = 0$. Also, we showed that the m out of n bootstrap is consistent for any constant μ. We also proposed the pre-test bootstrap which incorporates the pre-test for $\mu = 0$ and showed its consistency.

To compare the performances of the methods, we executed some Monte Carlo simulations. From the simulation results, it appears that the m out of n bootstrap method is generally well behaved for both $\mu = 0$ and $\mu \neq 0$ when m is large, but it can be beaten by the centered bootstrap method when $\mu = 0$ and by the conventional bootstrap method when μ is not zero. The pre-test bootstrap yields rather reliable results for all cases especially for the SR estimator. Also, the normal approximation is rather accurate when μ is far from 0. This may be because we assume X_is are distributed as a normal distribution. Because of the nature of the bootstraps, bootstrap methods will work when X_is are not distributed as a normal distribution and the normal approximation does not work very well. Furthermore, our simulation results show that the proposed methods are applicable in more general cases including the case of the general covariance matrix.

Thus, our simulation results showed that the validity of the pre-test bootstrap and the m out of n bootstrap. Though the order of the critical value in the pre-test are selected based on simulation results, the pre-test utilized in this paper are simple. Considering more sophisticated method or controlling the significance level of the pre-test based on given data may yield better performance of the bootstrap method. Also, applying these methods to the real data will be interesting field of research. However, these are beyond the scope of this research and remaining topic in the future.

Appendix A

Here, we derive the asymptotic distribution of the SR and the PSR estimators given in (3) and (4). From the central limit theorem, $\sqrt{n}(\overline{X}-\mu)\xrightarrow{d}Z$, where $Z\sim \mathrm{N}(0,\mathrm{I}_k)$. Thus, $\sqrt{n}\overline{X}\xrightarrow{d}Z$ when $\mu=0$, and $\sqrt{n}\overline{X}=O_p(n^{1/2})$ when $\mu\neq 0$. Therefore, we have

$$\frac{1}{n\overline{X}'\overline{X}}=\frac{1}{\sqrt{n}\overline{X}'\sqrt{n}\overline{X}}\xrightarrow{d}\frac{1}{Z'Z} \tag{20}$$

when $\mu=0$ and

$$\frac{1}{n\overline{X}'\overline{X}}\xrightarrow{\mathrm{pr}}0 \tag{21}$$

when $\mu\neq 0$. Since

$$\sqrt{n}(\widehat{\mu}_{\mathrm{SR}}-\mu)=\left(1-\frac{k-2}{n\overline{X}'\overline{X}}\right)\sqrt{n}(\overline{X}-\mu)-\frac{k-2}{n\overline{X}'\overline{X}}\sqrt{n}\mu, \tag{22}$$

and $\sqrt{n}\overline{X}=O_p(n^{1/2})$, we obtain the asymptotic distribution of $\widehat{\mu}_{\mathrm{SR}}$ given in (3). Similarly, since

$$\begin{aligned}\sqrt{n}(\widehat{\mu}_{\mathrm{PSR}}-\mu)&=\max\left[0,\ 1-\frac{k-2}{n\overline{X}'\overline{X}}\right]\sqrt{n}(\overline{X}-\mu)-\sqrt{n}\mu+\max\left[0,\ 1-\frac{k-2}{n\overline{X}'\overline{X}}\right]\sqrt{n}\mu\\&=\max\left[0,\ 1-\frac{k-2}{n\overline{X}'\overline{X}}\right]\sqrt{n}(\overline{X}-\mu)+\max\left[-1,\ -\frac{k-2}{n\overline{X}'\overline{X}}\right]\sqrt{n}\mu,\end{aligned} \tag{23}$$

we obtain the asymptotic distribution of $\widehat{\mu}_{\mathrm{PSR}}$ given in (4).

Appendix B

In this appendix, we derive the asymptotic distributions of the bootstrap estimators. We denote the bootstrap sample of size m (i.e., the random sample obtained by sampling with replacement from $\mathcal{X}=\{X_1,X_2,\ldots,X_n\}$) by $X_1^*,X_2^*,\ldots X_m^*$. By the central limit theorem and m/n tends to 0 as n tends to infinity, we have $\sqrt{m}(\overline{X}^*-\overline{X})\xrightarrow{d}Z^*$ given $\mathcal{X}=\{X_1,X_2,\ldots,X_n\}$, where $\overline{X}^*=\frac{1}{m}\sum_{i=1}^{m}X_i^*$ and $Z^*\sim\mathrm{N}(0,\mathrm{I}_k)$.

From the above result, when $m=n$ and $\mu=0$, we have $\sqrt{n}\overline{X}^*\overset{d}{\approx}Z^*+Z$. Note that Z is regarded as a realization from $\mathrm{N}(0,\mathrm{I}_k)$ since the bootstrap is executed given $\mathcal{X}$. Also, when $m=n$ and $\mu\neq 0$, $\sqrt{n}\overline{X}^*=O_p(n^{1/2})$. Since

$$\sqrt{n}(\widehat{\mu}^*_{\mathrm{SR}}-\overline{X})=\left(1-\frac{k-2}{n\overline{X}^{*\prime}\overline{X}^*}\right)\sqrt{n}(\overline{X}^*-\overline{X})-\frac{k-2}{n\overline{X}^{*\prime}\overline{X}^*}\sqrt{n}\overline{X}, \tag{24}$$

we obtain the asymptotic distribution of $\widehat{\mu}^*_{\mathrm{SR}}$ given in (7). By the similar transformation to (23), we can obtain the asymptotic distribution of $\widehat{\mu}^*_{\mathrm{PSR}}$ given in (8). Also, the asymptotic distributions of centered bootstrap estimators given in (11) and (12) can be obtained in similar ways.

As for the asymptotic distributions of the m out of n bootstrap estimators, since $\overline{X}=O_p(n^{1/2})$ when $\mu=0$, we have $\sqrt{m}\overline{X}\xrightarrow{\mathrm{pr}}0$ and $\sqrt{m}\overline{X}^*\xrightarrow{d}Z^*$. When $\mu\neq 0$, $\overline{X}\xrightarrow{\mathrm{pr}}\mu$ and $\sqrt{m}\overline{X}=O_p(m^{1/2})$. Using these results and the similar transformations to (24), we obtain the asymptotic distribution of the m out of n bootstrap SR estimator given in (15). The asymptotic distribution of the m out of n bootstrap PSR estimator can be obtained in a similar way.

References

1. A. J. Baranchik. A family of minimax estimators of the mean of a multivariate normal distribution. *The Annals of Mathematical Statistics*, 41:642–645, 1970.
https://doi.org/10.1214/aoms/1177697104
2. Rudolf Beran. Diagnosing bootstrap success. *Annals of the Institute of Statistical Mathematics*, 49:1–24, 1997.
https://doi.org/10.1023/a:1003114420352
3. Peter J Bickel, Friedrich Götze, and van Zwet. Resampling fewer than *n* observations: Gains, losses, and remedies for losses. *Statistica Sinica*, 7:1–31, 1997.
https://doi.org/10.1007/978-1-4614-1314-1_17
4. Peter J Bickel and Anat Sakov. On the choice of *m* in the *m* out of *n* bootstrap and confidence bounds for extrema. *Statistica Sinica*, 18:967–985, 2008.
5. David Brownstone. Bootstrapping improved estimators for linear regression models. *Journal of Econometrics*, 44(1):171–187, 1990.
https://doi.org/10.1016/0304-4076(90)90078-8
6. B. Efron. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26, 1979.
https://doi.org/10.1214/aos/1176344552
7. Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360, 2001.
https://doi.org/10.1198/016214501753382273
8. Patrik Guggenberger. The impact of a Hausman pretest on the asymptotic size of a hypothesis test. *Econometric Theory*, pages 369–382, 2010.
https://doi.org/10.1017/s0266466609100026
9. Bruce E Hansen. Efficient shrinkage in parametric models. *Journal of Econometrics*, 190(1):115–132, 2016a.
https://doi.org/10.1016/j.jeconom.2015.09.003
10. Bruce E Hansen. The risk of James-Stein and LASSO shrinkage. *Econometric Reviews*, 35(8-10):1456–1470, 2016b.
https://doi.org/10.1080/07474938.2015.1092799
11. Bruce E Hansen. Stein-Like 2SLS Estimator. *Econometric Reviews*, 36(6-9):840–852, 2017.
https://doi.org/10.1080/07474938.2017.1307579
12. Jerry A Hausman. Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pages 1251–1271, 1978.
https://doi.org/10.2307/1913827
13. W. James and Charles Stein. Estimation with quadratic loss. In Jerzy Neyman, editor, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 361–379. University of California Press, Berkeley, 1961.
14. Camilla Kazimi and David Brownstone. Bootstrap confidence bands for shrinkage estimators. *Journal of Econometrics*, 90:99–127, 1999.
https://doi.org/10.1016/s0304-4076(98)00037-2
15. Kazuhiro Ohtani. A comparison of the Stein-rule and positive-part Stein-rule estimators in a misspecified linear regression model. *Econometric Theory*, 9:668–679, 1993.
https://doi.org/10.1017/s0266466600008045
16. P. C. B. Phillips. The exact distribution of the Stein-rule estimator. *Journal of Econometrics*, 25:123–131, 1984.
https://doi.org/10.1016/0304-4076(84)90041-1
17. Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 197–206. University of California Press, Berkeley, 1956.
18. Robert Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
19. Aman Ullah. The approximate distribution function of the stein-rule estimator. *Economics Letters*, 10(3-4):305–308, 1982.
https://doi.org/10.1016/0165-1765(82)90070-2
20. Aman Ullah, Alan TK Wan, Huansha Wang, Xinyu Zhang, and Guohua Zou. A semiparametric generalized ridge estimator and link with model averaging. *Econometric Reviews*, 36:1–15, 2016.
https://doi.org/10.1080/07474938.2015.1114564
21. Hrishikesh D Vinod. Double bootstrap for shrinkage estimators. *Journal of Econometrics*, 68(2):287–302, 1995.
https://doi.org/10.1016/0304-4076(94)01639-h

22. Bei Wei, Stephen MS Lee, and Xiyuan Wu. Stochastically optimal bootstrap sample size for shrinkage-type statistics. *Statistics and Computing*, 26(1-2):249–262, 2016.
https://doi.org/10.1007/s11222-014-9493-x
23. Gang Yi. Estimating the variability of the Stein estimator by bootstrap. *Economics Letters*, 37(3):293–298, 1991.
https://doi.org/10.1016/0165-1765(91)90225-a
24. Hui Zou. The adaptive Lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429, 2006.
https://doi.org/10.1198/016214506000000735