Revista Colombiana de Estadística - Theoretical Statistics
January 2022, volume 45, issue 1, pp. 107–123
http://doi.org/10.15446/rce.v45n1.92037

Wavelet Shrinkage Generalized Bayes Estimation for Multivariate Normal Distribution Mean Vectors with unknown Covariance Matrix under Balanced-LINEX Loss

Contracción de la ondícula Estimación de Bayes generalizada para vectores medios de distribución normal multivariante con matriz de covarianza desconocida con pérdida de LINEX equilibrada

Hamid Karamikabir[a], Mahmud Afshari[b]

Department of Statistics, Faculty of Intelligent Systems Engineering and Data Science, Persian Gulf University, Bushehr, Iran

Abstract

In this paper, the generalized Bayes estimator of mean vector parameter for multivariate normal distribution with Unknown mean vector and covariance matrix is considered. This estimation is performed under the balanced-LINEX error loss function. The generalized Bayes estimator by using wavelet transformation is investigated. We also prove admissibility and minimaxity of shrinkage estimator and we present the simulation study and real data set for test validity of new estimator.

***Key words*:** admissibility; generalized bayes estimator; balanced-linex loss; minimaxity; multivariate normal distribution; soft wavelet shrinkage estimator.

Resumen

En este trabajo, se considera el estimador de Bayes generalizado del parámetro de vector medio para distribución normal multivariante con vector de media desconocido y matriz de covarianza. Esta estimación se realiza bajo la función de pérdida de error LINEX balanceada. Se investiga el estimador de Bayes generalizado mediante la transformación de ondículas. También probamos la admisibilidad y minimaxidad del estimador de contracción y presentamos el estudio de simulación y el conjunto de datos reales para comprobar la validez de la prueba del nuevo estimador.

***Palabras clave*:** admisibilidad; estimador de Bayes generalizado; estimador de contracción de ondas suaves; distribución normal multivariante; minimaxidad; pérdida de LINEX equilibrada.

[a]Ph.D. E-mail: h_karamikabir@pgu.ac.ir
[b]Ph.D. E-mail: afshar@pgu.ac.ir

1. Introduction

Many univariate tests and confidence intervals are based on the univariate normal distribution. Similarly, the majority of multivariate procedures have the multivariate normal distribution as their underpinning. For more information about this issue refer to Rencher & Christensen (2012).

Given the importance of the multivariate normal distribution, estimating the parameters of this distribution is very important. In this paper, our goal is to estimate the mean vector of this distribution. We suppose that the random vector $X = (X_1, \ldots, X_p)$ has a multivariate normal distribution with unknown mean vector $\theta = (\theta_1, \ldots, \theta_p)$ and unknown covariance matrix Σ ($X \sim N_p(\theta, \Sigma)$). For this purpose, using the generalized Bayes estimator, we estimated mean vector under a combination of balanced loss function and asymmetric linear exponential (LINEX) loss function.

Balanced loss functions and their role in estimation have captured the interest of many researchers. The balanced loss function was introduced by Zellner (2009) to reflect two criteria: goodness of fit and precision of estimation. In Zellner's framework, the target estimator was least-squares, but such a target can be viewed more broadly (e.g., Jozani, Marchand & Parsian 2006, 2014). For more details about the use of this loss, we refer to Jozani et al. (2012), Cao & He (2017), Zinodiny et al. (2017), Karamikabir & Arashi (2018), Karamikabir et al. (2020), Marchand & Strawderman (2020), and Karamikabir & Afshari (2020, 2021) to mention a few.

Suppose that X is a random vector having a multivariate normal distribution with mean vector parameter θ. The balanced-type loss function is defined as follows:

$$L^*_{\omega,\delta_0}(\theta,\delta) = \omega\rho\left(\delta_0(X), \delta(X)\right) + (1-\omega)\rho\left(\theta, \delta(X)\right), \qquad 0 \leq \omega < 1. \quad (1)$$

where δ_0 is a target estimator of θ obtained for instance using the criterion of maximum likelihood, least-squares, unbiasedness etc. $\rho(\cdot)$ is an arbitrary multivariate loss function and $\delta(X)$ is an estimator of p-vector parameter θ based on the random vector X.

In the balanced-type loss function (1), with the change of multivariate loss function $\rho(\cdot)$, we can define different types of the balanced-type loss functions. If we consider multivariate LINEX loss function $\rho(l,\delta) = b\Big(e^{a^T(\delta-l)} - a^T(\delta-l) - 1\Big)$ with the scale parameter b and the p-vector shape parameter $a = (a_1, \ldots, a_p)^T$, then we have the balanced-LINEX loss function as follows:

$$\begin{aligned} L_{\omega,\delta_0}(\theta,\delta) &= b\omega\left(e^{a^T(\delta(X)-\delta_0(X))} - a^T(\delta(X)-\delta_0(X)) - 1\right) \\ &\quad + b(1-\omega)\left(e^{a^T(\delta(X)-\theta)} - a^T(\delta(X)-\theta) - 1\right), \quad 0 \leq \omega < 1. \quad (2) \end{aligned}$$

If $\omega = 0$, then the balanced-LINEX loss function becomes the basic case of LINEX loss function. For more information about this loss function see Jozani et al. (2012).

Shrinking and truncating the data directly or the coefficients in their Fourier series expansions is an old technique in signal and image processing. For non-local bases, such as trigonometric, shrinking the coefficients can affect the global shape of the reconstructed function and introduce unwanted artifacts. In the context of function estimation by wavelets, the shrinkage has an additional feature; it is connected with smoothing (denoising) because the measures of smoothness of a function depend on the magnitudes of its wavelet coefficients (Vidakovic, 2009). Wavelet methods are usually employed as a form of nonparametric regression, and the techniques take on many names such as wavelet shrinkage, curve estimation, or wavelet regression. Two different kinds of threshold in denoising is the hard threshold and the soft threshold. Donoho & Johnstone (1994) define the hard and soft thresholding functions. Given a wavelet coefficient X and a threshold value $\lambda > 0$, the hard threshold value is given by

$$\delta^{hard}(X) = XI(|X| \geq \lambda),$$

and the soft thresholding wavelet shrinkage estimation is given by

$$\delta^{soft}(X) = sign(X)(|X| - \lambda)I(|X| \geq \lambda). \tag{3}$$

where $I(\cdot)$ is an indicator function.

We try to make a connection between the generalized Bayes estimator and the wavelet shrinkage estimator. For this purpose, we will find a threshold value for the soft thresholding wavelet shrinkage estimator using the generalized Bayes estimator.

In this paper, we also generalized the paper of Karamikabir & Afshari (2019) by changing loss function and the covariance matrix and Karamikabir & Afshari (2020) by changing the diagonal covariance matrix $\sigma^2 I_p$ to the unknown covariance matrix Σ. Recently, the problem of estimating a mean vector parameter has received several new developments. For example, we can refer to Pal et al. (2007), Jiang & Zhang (2009), Tsukuma & Kubokawa (2015), Fourdrinier & Strawderman (2015), Joly & Oliveira (2017), Karamikabir & Arashi (2018) and Karamikabir & Afshari (2020).

Finally, we present a method to select the threshold value in wavelet regularization. For this purpose, the threshold value is selected using the generalized Bayes estimator and the method described in Section 3.

The paper is outlined as follows. In Section 2, we find the generalized Bayes estimator when $X \sim N_p(\theta, \Sigma)$ under a balanced-LINEX loss function. In Section 3, we discuss the shrinkage Wavelet generalized Bayes estimation and threshold value, and in Section 4 the numerical performance of the proposed estimator using a simulation study. In Section 5 we investigate a real example, and in Section 6 concludes the paper.

2. Main Result

In this section, we investigate the point estimation of the mean vector θ when the covariance matrix Σ is unknown. For this purpose, we suppose that

$X \sim N_p(\theta, \Sigma)$ and we find the generalized Bayes estimator for θ with respect to the improper prior $\pi(\theta) = 1$ under the balanced-LINEX loss function. Also suppose that $X|\theta \sim N_p(\theta, \Sigma)$ and $\delta(X) = (\delta(X_1), \ldots, \delta(X_p))^T$ be an estimator for θ. In this regard, we need the following theorem.

Theorem 1 (Rudin 1976, Chapter 7, pp. 148)**.** *Suppose* $\lim_{n\to\infty} f_n(x) = f(x)$ *where* $x \in H$*. Put* $M_n = \sup_{x\in H} |f_n(x) - f(x)|$*. Then* $f_n \to f$ *uniformly on* H *if and only if* $M_n \to 0$ *as* $n \to \infty$*.*

Theorem 2 (Rudin 1976, Chapter 7, pp. 167)**.** *Suppose* $f_n(x)$ *is Riemann integrable on* $[a, b]$*, for* $n = 1, 2, \ldots,$ *and suppose* $f_n(x) \to f(x)$ *uniformly on* $[a, b]$*. Then* $f(x)$ *is Riemann integrable on* $[a, b]$*, and*

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

Now, in the following theorem, we find the generalized Bayes estimator for θ under the balanced-LINEX loss function in (2).

Theorem 3. *Suppose that* $X \sim N_p(\theta, \Sigma)$*,* $\theta \in \mathbb{R}^p$*, under the balanced-LINEX loss function, the generalized Bayes estimator for* θ *with respect to the improper prior* $\pi(\theta) = 1$*, is the following:*

$$\delta^*(X) = \frac{-a}{\|a\|^2} \log\left[\omega \exp\left(-a^T \delta_0(X)\right) + (1-\omega)\exp\left(-a^T X + \frac{1}{2}a^T \Sigma a\right)\right].$$

Proof. For $\pi(\theta) = 1$, the posterior distribution is $\pi(\theta|X) \sim N_p(X, \Sigma)$. It is easy to check that the posterior loss function of an arbitrary estimator $\delta(X)$ is given by

$$\begin{aligned} r^*\left(\pi(\theta|X), \delta(X)\right) &= \int_\Theta L_{\omega,\delta_0}(\theta, \delta)\pi(\theta|X)d\theta \\ &= E\Big(b\omega\left(e^{a^T(\delta(X)-\delta_0(X))} - a^T(\delta(X) - \delta_0(X)) - 1\right) \\ &\quad + b(1-\omega)\left(e^{a^T(\delta(X)-\theta)} - a^T(\delta(X) - \theta) - 1\right)\Big|X\Big). \end{aligned}$$

The generalized Bayes estimator is obtained by the following:

$$\begin{aligned} \frac{\partial r^*\left(\pi(\theta|X), \delta(X)\right)}{\partial \delta(X)} &= bE\Big(\omega a^T e^{a^T(\delta(X)-\delta_0(X))} + (1-\omega)a^T e^{a^T(\delta(X)-\theta)} - a^T\Big|X\Big) \\ &= b\left(\omega a e^{a^T(\delta(X)-\delta_0(X))} + (1-\omega)a e^{a^T\delta(X)} M_{\theta|X}(-a^T) - a\right) \\ &= 0, \end{aligned}$$

then the generalized Bayes estimator as follows:

$$\delta^*(X) = \frac{-a}{\|a\|^2} \log\left[\omega \exp\left(-a^T \delta_0(X)\right) + (1-\omega)\exp\left(-a^T X + \frac{1}{2}a^T \Sigma a\right)\right].$$

□

In Theorem 3, we obtained the generalized Bayes estimator $\delta^*(X)$ with respect to the target estimator $\delta_0(X)$. By changing the target estimator $\delta_0(X)$ and the covariance matrix Σ, the generalized Bayes estimator $\delta^*(X)$ takes on different types. In this regard, we have the following corollaries by Theorem 3:

Corollary 1. *Suppose that $X \sim N_p(\theta, \sigma^2 I_p)$, then the generalized Bayes estimator is*

$$\delta^*(X) \;=\; \frac{-a}{\|a\|^2}\log\left[\omega\exp\left(-a^T\delta_0(X)\right)+(1-\omega)\exp\left(-a^TX+\frac{1}{2}\sigma^2\|a\|^2\right)\right].$$

Also, if that $\omega = 0$, then the generalized Bayes estimator is

$$\delta^*(X) \;=\; X-\frac{\sigma^2}{2}a.$$

Corollary 2. *Suppose that $X \sim N_p(\theta, \Sigma)$ and $\delta_0(X) = X$, then the generalized Bayes estimator is*

$$\delta^*(X) = X-\frac{a}{\|a\|^2}\log\left(\omega+(1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right). \tag{4}$$

Also, if that $\Sigma = \sigma^2 I_p$, then the generalized Bayes estimator is

$$\delta^*(X) = X-\frac{a}{\|a\|^2}\log\left(\omega+(1-\omega)\exp\left(\frac{\sigma^2}{2}\|a\|^2\right)\right). \tag{5}$$

The following proposition shows that the results of Huang (2002), Torehzadeh & Arashi (2014) and Karamikabir & Afshari (2019) are only a special case of Theorem 3.

Proposition 1. *The special case of Theorem 3 is the following:*

1. *Suppose that $X \sim N(\theta, \sigma^2 I_p)$. The generalized Bayes estimator for $\omega = 0$, $a^T = (0, \ldots, a_i, \ldots, 0)$ and $b = \frac{1}{n}$ is as follows:*

$$\delta^*(X_i) = X_i - \frac{a_i\sigma^2}{2}.$$

See Huang (Huang, 2002).

2. *Suppose that $X \sim SMN(\theta, \sigma^2, G)$ (the covariance matrix of multivariate normal distribution). The generalized Bayes estimator for $\omega = 0$, $a^T = (0, \ldots, a_i, \ldots, 0)$ and $b = \frac{1}{n}$ is as follows:*

$$\delta^*(X_i) = X_i - \frac{\ln\alpha(a_i^2, \sigma^2)}{a_i},$$

where $\alpha(a_i^2, \sigma^2) = \int_0^\infty e^{\frac{a_i^2\sigma^2}{2t}}dG(t)$. See Torehzadeh & Arashi (2014).

3. *Suppose that* $X \sim N_p(\theta, \Sigma)$*, The generalized Bayes estimator for* $\omega = 0$*,* $a^T = (0, \ldots, a_i, \ldots, 0)$ *and* $b = \frac{1}{n}$ *is as follows:*

$$\delta^*(X_i) = X_i - \frac{a_i \sigma_{ii}}{2}.$$

See Karamikabir & Afshari (2019).

Now, we want to find minimax and admissible estimator based on general Bayes estimator. Suppose that $X \sim N_p(\theta, \Sigma)$ where θ and Σ is unknown. Also suppose that $\pi(\theta)$ be an arbitrary proper prior, under the balanced-LINEX loss function, the Bayes risk of the estimator $\delta^*(X)$ is given by

$$\begin{aligned} r(\pi(\theta), \delta^*(X)) &= \int_\Theta R(\theta, \delta^*(X))\pi(\theta) d\theta \\ &= \int_\Theta \pi(\theta) \int_X b\omega \left(e^{a^T(\delta(x)-\delta_0(x))} - a^T(\delta(x) - \delta_0(x)) - 1 \right) \\ &\quad + b(1-\omega) \left(e^{a^T(\delta(x)-\theta)} - a^T(\delta(x) - \theta) - 1 \right) dx \, d\theta. \end{aligned} \tag{6}$$

The target estimator obtained using the criterion of the maximum likelihood, least-squares, unbiasedness and etc. We have selected target estimator as the maximum likelihood $\delta_0(X) = X$ for the balanced-LINEX loss function (see Corollary 2).

In the following theorem, we want to find the the Bayes risk of the estimator $\delta^*(X)$ by using the equation (6). For this purpose, we first obtain the risk $R(\theta, \delta^*(X))$ and then integrate it with respect to θ. If the the Bayes risk $r(\pi(\theta), \delta^*(X))$ is constant value, the generalized Bayes estimator $\delta^*(X)$ estimator is Minimax.

Theorem 4. *Suppose that* $X \sim N_p(\theta, \Sigma)$*,* $\pi(\theta)$ *be an arbitrary proper prior and* $\delta_0(X) = X$*. Under the balanced-LINEX loss function in* (2) *the Bayes risk of the estimator* $\delta^*(X)$ *in* (4) *is given by*

$$r(\pi(\theta), \delta^*(X)) = b \log \left(\omega + (1-\omega) \exp \left(\frac{1}{2} a^T \Sigma a \right) \right). \tag{7}$$

Also the generalized Bayes estimator is minimax.

Proof. By using Corollary 2, we have the generalized Bayes estimator $\delta^*(X)$ in (4). In this case, the risk function calculated as follows.

$$\begin{aligned} R(\theta, \delta^*(X)) &= E\left[L_{\omega,\delta_0}(\theta, \delta(X))\right] \\ &= E\Bigg\{ b\omega \Bigg[\exp \left(a^T \left(X - \frac{a}{\|a\|^2} \log \left(\omega + (1-\omega) e^{\frac{1}{2} a^T \Sigma a} \right) - X \right) \right) \\ &\quad - a^T \left(X - \frac{a}{\|a\|^2} \log \left(\omega + (1-\omega) e^{\frac{1}{2} a^T \Sigma a} \right) - X \right) - 1 \Bigg] \\ &\quad + b(1-\omega) \Bigg[\exp \left(a^T \left(X - \frac{a}{\|a\|^2} \log \left(\omega + (1-\omega) e^{\frac{1}{2} a^T \Sigma a} \right) - \theta \right) \right) \end{aligned}$$

$$
\begin{aligned}
& -a^T\left(X-\frac{a}{\|a\|^2}\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)-\theta\right)-1\Bigg]\Bigg\} \\
= \; & b\omega\left[\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)^{-1}+\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)-1\right] \\
& +b(1-\omega)\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)^{-1}e^{-a^T\theta}M_{X|\theta}(a^T) \\
& +b(1-\omega)\left[E\left(-a^TX|\theta\right)+\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)+a^T\theta-1\right] \\
= \; & b\omega\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)^{-1}+b\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)-b \\
& +b(1-\omega)\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right)^{-1}e^{\frac{1}{2}a^T\Sigma a} \\
= \; & b\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right).
\end{aligned} \tag{8}
$$

By using the risk $R(\theta,\delta^*(X))$ in (8), we have the following Bayes risk

$$
\begin{aligned}
r(\pi(\theta),\delta^*(X)) &= \int_\Theta R(\theta,\delta^*(X))\pi(\theta)d\theta \\
&= b\log\left(\omega+(1-\omega)e^{\frac{1}{2}a^T\Sigma a}\right).
\end{aligned}
$$

Since Bayes risk $r(\pi(\theta),\delta^*(X))$ is a constant value, so $\delta^*(X)$ is a minimax estimator. □

Corollary 3. *In Theorem 4, suppose that $X \sim N_p(\theta,\sigma^2 I_p)$, then the Bayes risk of the estimator $\delta^*(X)$ in (5) is given*

$$
r(\pi(\theta),\delta^*(X)) = b\log\left(\omega+(1-\omega)\exp\left(\frac{\sigma^2}{2}\|a\|^2\right)\right).
$$

Also, if $\omega=0$, then the Bayes risk of the estimator $\delta^*(X)$ in (4) is given by

$$
r(\pi(\theta),\delta^*(X)) = \frac{b}{2}a^T\Sigma a.
$$

And the Bayes risk of the estimator $\delta^*(X)$ in (5) is given by

$$
r(\pi(\theta),\delta^*(X)) = \frac{b\;\sigma^2}{2}\|a\|^2.
$$

In the following Lemma, we find the posterior distribution and generalized Bayes estimators a multivariate normal distribution $N_p(\theta,\Sigma)$ with conjugate prior distribution $N_p(0,\varrho_k^2\Sigma)$.

Lemma 1. *Suppose that $X|\theta \sim N_p(\theta,\Sigma)$ and the prior distribution $\pi_k(\theta) = N_p(0,\varrho_k^2\Sigma)$, then, we have the following results.*

- *The posterior distribution $\pi(\theta|X)$ is $N_p\left(\frac{\varrho_k^2}{\varrho_k^2+1}X, \frac{\varrho_k^2}{\varrho_k^2+1}\Sigma\right)$.*

- *Under the balanced-LINEX loss function in (2), when $\delta_0(X) = X$, the Bayes estimator for θ is:*

$$\delta^{\pi_k}(X) = X - \frac{a}{\|a\|^2} \log\left[\omega + (1-\omega)\exp\left(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\right)\right]. \quad (9)$$

Proof. The posterior distribution $\pi(\theta|X)$ is obtained as follows.

$$\begin{aligned}
\pi(\theta|X) &\propto f(\omega|\theta)\ \pi(\theta) \\
&\propto \exp\left[-\frac{1}{2}\left(x^T\Sigma^{-1}x + \theta^T\Sigma^{-1}\theta - 2\theta^T\Sigma^{-1}x + \frac{1}{\varrho_k^2}\theta^T\Sigma^{-1}\theta\right)\right] \\
&\propto \exp\left[-\frac{1}{2}\theta^T\left(\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1}\right)\theta + \theta^T\Sigma^{-1}x\right] \\
&= \exp\left\{-\frac{1}{2}\left(\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1}\right)\left[\theta^T\theta - 2\theta^T\Sigma^{-1}(\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1})^{-1}x\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\Big[\left(\theta - (\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1})^{-1}\Sigma^{-1}x\right)^T\left(\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1}\right)^{-1}\right. \\
&\quad \left.\times\left(\theta - (\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1})^{-1}\Sigma^{-1}x\right)\Big]\right\} \\
&\quad \sim N_p\left((\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1})^{-1}\Sigma^{-1}X, (\Sigma^{-1} + \frac{1}{\varrho_k^2}\Sigma^{-1})^{-1}\right) \\
&\quad = N_p\left(\frac{\varrho_k^2}{\varrho_k^2+1}X, \frac{\varrho_k^2}{\varrho_k^2+1}\Sigma\right).
\end{aligned}$$

Now similar to Theorem 3, a generalized Bayes estimator is obtained by the following:

$$\begin{aligned}
\frac{\partial r^*\left(\pi(\theta|X), \delta(X)\right)}{\partial\delta(X)} &= bE\Big(\omega a^T e^{a^T(\delta(X)-\delta_0(X))} + (1-\omega)a^T e^{a^T(\delta(X)-\theta)} - a^T\Big|X\Big) \\
&= b\left(\omega a^T e^{a^T(\delta(X)-\delta_0(X))} + (1-\omega)a^T e^{a^T\delta(X)}M_{\theta|X}(-a^T) - a^T\right) \\
&= 0,
\end{aligned}$$

then we have the following Bayes estimator with respect to the target estimator $\delta_0(X)$.

$$\delta^{\pi_k}(X) = \frac{-a}{\|a\|^2}\log\left(\omega e^{-a^T\delta_0(X)} + (1-\omega)\exp\left(-\frac{\varrho_k^2}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T\Sigma a\right)\right).$$

Finally by replacing $\delta_0(X) = X$ in the Bayes estimator $\delta^{\pi_k}(X)$, we have the following Bayes estimator.

$$\delta^{\pi_k}(X) = X - \frac{a}{\|a\|^2}\log\left(\omega + (1-\omega)\exp\left(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T\Sigma a\right)\right).$$

□

Now, in the following theorem, we prove that the generalized Bayes estimators $\delta^*(X)$ is admissible and minimax under the balanced-LINEX loss function.

Theorem 5. *Suppose that $X \sim N_p(\theta, \Sigma)$ and $\delta_0(X) = X$, then under the balanced-LINEX loss function, $\delta^*(X)$ in (4) is admissible and a minimax estimator.*

Proof. We know $R(\theta, \delta)$ is continuous in θ for any δ. Suppose that δ^* is not admissible. Then, there exists an estimator δ such that $R(\theta, \delta) < R(\theta, \delta^*)$, with strict inequality for some θ, say θ_0. Since $R(\theta, \delta)$ and $R(\theta, \delta^*)$ are continuous in θ, there exist strictly positive constants c_1 and c_2 such that

$$R(\theta, \delta) < R(\theta, \delta^*) - c_1 \quad for \quad \theta \in [\theta : |\theta - \theta_0| < c_2],$$

Consider a sequence of priors $\pi_k(\theta) = N_p(0, \varrho_k^2 \Sigma)$, with $\lim_{k\to\infty} \varrho_k^2 = \infty$, and uniformly, $\lim_{k\to\infty} \pi_k(\theta) \to \pi(\theta)$, where $\pi(\theta) < \infty$ is a proper distribution. Using the technique of minimizing posterior expected loss, under the balanced-LINEX loss function by using Theorem 1, we have the Bayes estimator $\delta^{\pi_k}(X)$ in (9). By using the balanced-LINEX loss function $L_{\omega,\delta_0}(\theta, \delta(X))$ in (2), the risk $R(\theta, \delta^{\pi_k}(X))$ as follows:

$$\begin{aligned}
& bE \quad \Bigg\{\omega \exp\Bigg\{-\log\Bigg[\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg]\Bigg\} \\
& \quad +\omega \log\Bigg[\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg] - 1 \\
& \quad +(1-\omega)\exp\Bigg\{a^T X - a^T \log\Bigg[\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X \\
& \quad + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg] - a^T\theta\Bigg\} - (1-\omega)\Bigg[a^T X - a^T\theta \\
& \quad - \log\Bigg[\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg]\Bigg]\Bigg\} \\
= & \quad bE\Bigg[\omega\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg)^{-1} \\
& \quad + \log\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg) - 1\Bigg] \\
& \quad +b(1-\omega)M_{X|\theta}(a^T X)e^{-a^T\theta} \\
& \quad E\Bigg[\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg)^{-1}\Bigg] \\
& \quad -b(1-\omega)E\left(a^T X|\theta\right) + b(1-\omega)a^T\theta \\
= & \quad bE\Bigg[\omega\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg)^{-1} \\
& \quad + \log\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg) - 1 \\
& \quad +(1-\omega)\Bigg(\omega + (1-\omega)\exp\Bigg(\frac{1}{\varrho_k^2+1}a^T X + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T \Sigma a\Bigg)\Bigg)^{-1}
\end{aligned}$$

$$\times \exp\left(\frac{1}{2}a^T\Sigma a\right)\Bigg]. \tag{10}$$

By using Theorem 1 for equation (10), we can obtain the values of risk as follows.

$$\begin{aligned} r(\pi_k(\theta), \delta^{\pi_k}(X)) &= \int_\Theta R(\theta, \delta^{\pi_k}(x))\pi(\theta)d\theta \\ &= \int_\Theta \int_\chi g_k(x)dx\ d\theta. \end{aligned}$$

where

$$\begin{aligned} g_k(x) &= b\omega\left(\omega + (1-\omega)\exp\left(\frac{1}{\varrho_k^2+1}a^TX + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T\Sigma a\right)\right)^{-1} \\ &+ \log\left(\omega + (1-\omega)\exp\left(\frac{1}{\varrho_k^2+1}a^TX + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T\Sigma a\right)\right) - 1 \\ &+(1-\omega)\left(\omega + (1-\omega)\exp\left(\frac{1}{\varrho_k^2+1}a^TX + \frac{\varrho_k^2}{2(\varrho_k^2+1)}a^T\Sigma a\right)\right)^{-1} \\ &\times \exp\left(\frac{1}{2}a^T\Sigma a\right). \end{aligned}$$

As a result

$$\begin{aligned} \lim_{k\to\infty} g_k(x) &= b\omega\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right)^{-1} \\ &+b\log\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right) - b \\ &+b(1-\omega)\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right)^{-1}\exp\left(\frac{1}{2}a^T\Sigma a\right) \\ &= b\log\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right) = g(x). \end{aligned}$$

Again, we put that $M_k = \sup_{x\in\mathbb{R}^p}|g_k(x) - g(x)|$, then we can write

$$\begin{aligned} \lim_{k\to\infty}\sup_{x\in\mathbb{R}^p}|g_k(x)-g(x)| &= b\log\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right) \\ &-b\log\left(\omega + (1-\omega)\exp\left(\frac{1}{2}a^T\Sigma a\right)\right) = 0. \end{aligned}$$

By Theorem 1, since $\lim_{k\to\infty} M_k = 0$, then $\lim_{k\to\infty} g_k(x) = g(x)$ uniformly on $\mathbb{R}^p$. As a result by Theorem 2, we can write

$$\lim_{k\to\infty} r(\pi_k(\theta), \delta^{\pi_k}(X)) = b\log\left(\omega + (1-\omega)\psi\left(\sigma^2 a^T a\right)\right).$$

Now, according to $r(\pi(\theta), \delta^*(X))$ in equation (7) we have

$$\lim_{k\to\infty}\{r(\pi(\theta), \delta^*(X)) - r(\pi_k(\theta), \delta^{\pi_k}(X))\} = 0. \tag{11}$$

Let $c_3 = \lim_{k\to\infty} \inf \int_{|\theta-\theta_0|<c_2} \pi_k(\theta)d\theta$. Since $\lim_{k\to\infty} \varrho_k^2 = \infty$, then $c_3 > 0$. Therefore, for k is large enough

$$\begin{aligned} r(\pi_k, \delta^*) - r(\pi_k, \delta^{\pi_k}) &\geq r(\pi_k, \delta^*) - r(\pi_k, \delta) \\ &= \int_{R^p} \left(R\left(\theta, \delta^*\right) - R\left(\theta, \delta\right)\right) \pi_k(\theta) d\theta > c_1 c_2 > 0. \end{aligned}$$

This contradicts with equation (11). As a result, $\delta^*(X)$ is an admissible estimator. The minimaxity of $\delta^*(X)$ follows from its admissibility and the constant risk phenomenon (7). □

3. Shrinkage Wavelet Generalized Bayes Estimation

In this section, the goal is to find a particular type of the soft wavelet estimator using the generalized Bayes estimator. In the issue, Huang (2002) investigated the shrinkage wavelet estimation problem in the multivariate normal by diagonal covariance matrix $\sigma^2 I_p$ and Torehzadeh & Arashi (2014) extended his result for a scale mixture of multivariate normal distributions. Finally Karamikabir & Afshari (2019) investigated the shrinkage wavelet estimation problem in the class of elliptically distribution, in LINEX loss function.

Consider the following model:

$$X = \theta + \varepsilon,$$

where $X = (X_1, \ldots, X_p)^T$ are the $p \times 1$ random vector, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)^T$ are independent identical distribution $N_p(0, \Sigma)$, and $\theta = (\theta_1, \ldots, \theta_p)^T$ are the p-vector mean vectors. Again, suppose that $X|\theta \sim N_p(\theta, \Sigma)$ and $\delta(X) = (\delta(X_1), \ldots, \delta(X_p))^T$ be an estimator for θ.

For deonising or shrinkage coefficients, one of the most important concepts in wavelets and deionising is using thresholds. Shrinkage of the empirical wavelet coefficients works best in problems where the underlying set of the true coefficients of f is *sparse* and the remaining few large coefficients explain most of the functional form in f. By shrinking, the empirical coefficients towards zero, the smaller ones which contain primarily noise may be reduced to negligible levels, hence denoising the signal.

Let $Y_1, \ldots, Y_n$ are observed data from model,

$$Y = f(Z) + \eta,$$

where the $\{\eta_i\}$ is some noise and $\{Z_i\}$ is some points from domain of f. Typically n is an integer power of 2.

Note that the observations are sampled from distribution f but with some noise and we are interested to remove noises. To achieve this aim, observations or noisy data are converted to wavelet coefficients. Denoised coefficients are returned to the Y domain by the inverse discrete wavelet transformation.

In this section, we suppose that $X \sim N_p(\theta, \sigma^2 I_p)$. In this regard, we use the condition of corollary 2, under the balanced-LINEX loss function with $\delta_0(X) = X$. Consider the generalized Bayes estimator in (4). For $a^T = (0, \ldots, a_i, \ldots, 0)$ (ith element is a_i), the ith element of $\delta^*(X)$ is

$$\delta^*(X_i) = X_i - \frac{1}{a_i} \log \left(\omega + (1-\omega) \exp\left(\frac{1}{2} a_i^2 \sigma^2 \right) \right).$$

We again suppose that $a^T = (0, \ldots, a_i, \ldots, 0)$ in the balanced-LINEX loss function. We consider specifically a_i values depending on signs of θ_i's

$$a_i = \begin{cases} c & for,\ \theta_i > 0, \\ -c & for,\ \theta_i < 0, \end{cases} \quad i = 1, \ldots n.$$

where $c > 0$ is some constant. Such an error criterion discourages estimators from over-estimation in magnitude (i.e., in absolute value) and results in shrinkage estimation towards zero. In other words, we can be considered this issue as a regularization problem that regularizes or shrinks the wavelet coefficient estimates towards zero.

Under such a loss criterion the generalized Bayes estimator in (4) is given by $\delta^*(X_i) = X_i - sign(\theta_i)\lambda_i$, where $\lambda_i = \frac{1}{c} \log\left(\omega + (1-\omega)\exp\left(\frac{\sigma_{ii}}{2} c^2\right)\right)$.

The wavelet estimation problem can be treated via the estimation of the mean vector θ from a elliptical distribution $X|\theta \sim N_p(\theta, \Sigma)$. Often the signs of parameters θ_i's are not known. A natural approach is to use sign(X_i) to estimate sign(θ_i) and make truncation at zero. In conclusion, we have the empirical version of δ^* in the following Proposition.

Proposition 2. *According to $\delta^{soft}(X_i)$ in (3), by choosing the threshold value of $\lambda_i = \frac{1}{c} \log\left(\omega + (1-\omega)\exp\left(\frac{\sigma_{ii}}{2} c^2\right)\right)$, the soft wavelet shrinkage estimator for θ by using $\delta^{soft}(X_i)$ can be obtained as follows.*

$$\delta^{soft}(X_i) = \begin{cases} (X_i - \lambda_i) \vee 0 & for,\ X_i \geq 0, \\ (X_i + \lambda_i) \wedge 0 & for,\ X_i < 0, \end{cases} = sign(X_i)(|X_i| - \lambda_i)_+.$$

4. Simulation

In this section, we checked theoretical outcomes with the numerical computation and simulation to investigate the performance of the soft wavelet shrinkage estimator in Section 3.

We compare the new threshold method to the three commonly used shrinkage strategies, i.e, hard and soft thresholding with the universal threshold and false discovery rate (FDR). To assess the performance, we calculated the average mean squared error (AMSE) from the $m = 1000$ simulations. The value of AMSE is obtained as follows:

$$m^{-1} N^{-1} \sum_{j=1}^{m} \sum_{i=1}^{N} \left(f(x_i) - \hat{f}(x_{i,j}) \right)^2,$$

where $f(x_i)$ is the true signal and $\hat{f}(x_{i,j})$ is the estimate of the function from simulation j. In general, lower values of AMSE represent the accuracy of the estimate. We suppose that $X \sim N_2(\theta, \Sigma)$ where $\theta = (0, 0)$ and $\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$. Tables 1 and 2 represent the AMSE with respect to c and ω for wavelet estimator based on the hard and soft universal threshold, FDR threshold and new threshold for the $X \sim N_2(\theta, \Sigma)$. As shown in Tables 1 and 2, the AMSE amount obtained in the new method is lower than that of the methods. Also, by increasing of the value of ω, the estimation AMSE increases and by increasing the N and c, the AMSE of the all method decreases.

Table 1: AMSE for hard and soft universal, FDR and new threshold for $c = 25$.

N	New threshold			Universal	Universal	FDR
	$\omega = 0.2$	$\omega = 0.5$	$\omega = 0.8$	Soft	Hard	
128	0.01615664	0.01616252	0.01617411	0.04245191	0.22163175	0.11313481
256	0.00931865	0.00932817	0.00934689	0.03439703	0.19397322	0.11066790

Table 2: AMSE for hard and soft universal, FDR and new threshold for $c = 35$.

N	New threshold			Universal	Universal	FDR
	$\omega = 0.2$	$\omega = 0.5$	$\omega = 0.8$	Soft	Hard	
128	0.01550174	0.01550226	0.01550329	0.04245191	0.22163175	0.11313481
256	0.00798796	0.00798963	0.00799290	0.03439703	0.19397322	0.11066790

In general, as the amount of ω increases, the risk increases. The reason for this is the value of threshold (λ). As the amount of ω increases, the amount of λ decreases. Also, by increasing the c or σ_{ii}, the AMSE decreases. Because the amount of c or σ_{ii} increases, the amount of λ increases.

Now, we checked theoretical outcomes with the numerical computation and simulation to investigate the performance of the soft wavelet shrinkage estimator and generalized Bayes estimator. All calculations in this section are done using **R** software.

To investigate the risk of estimators, a Monte Carlo simulation study was performed to compare the risk values estimators for the $N_8(\theta, \Sigma)$ where Σ is randomly generated using Wishart distribution and θ is selected as $\left(\sqrt{k}, 0, \ldots, 0\right)$ and $k = 0, 0.1, 0.2, \ldots, 10$. In this case, $\|\theta\| = \theta^T\theta = \sum_{i=1}^{p} {\theta_i}^2 = k$. These risk values have been obtained using the 1000 Monte Carlo simulation replications and plotted in Figure 1 for $p = 8$, $c = 25$, different values of ω, the soft wavelet shrinkage estimator and generalized Bayes estimator.

In Figure 1 the soft wavelet shrinkage estimator risk curve is lower that of the generalized Bayes estimator, i.e., the soft wavelet shrinkage estimator dominates the generalized Bayes estimator. As the value of ω increases, the superiority of the soft wavelet shrinkage estimator over the generalized Bayes estimator is increases.

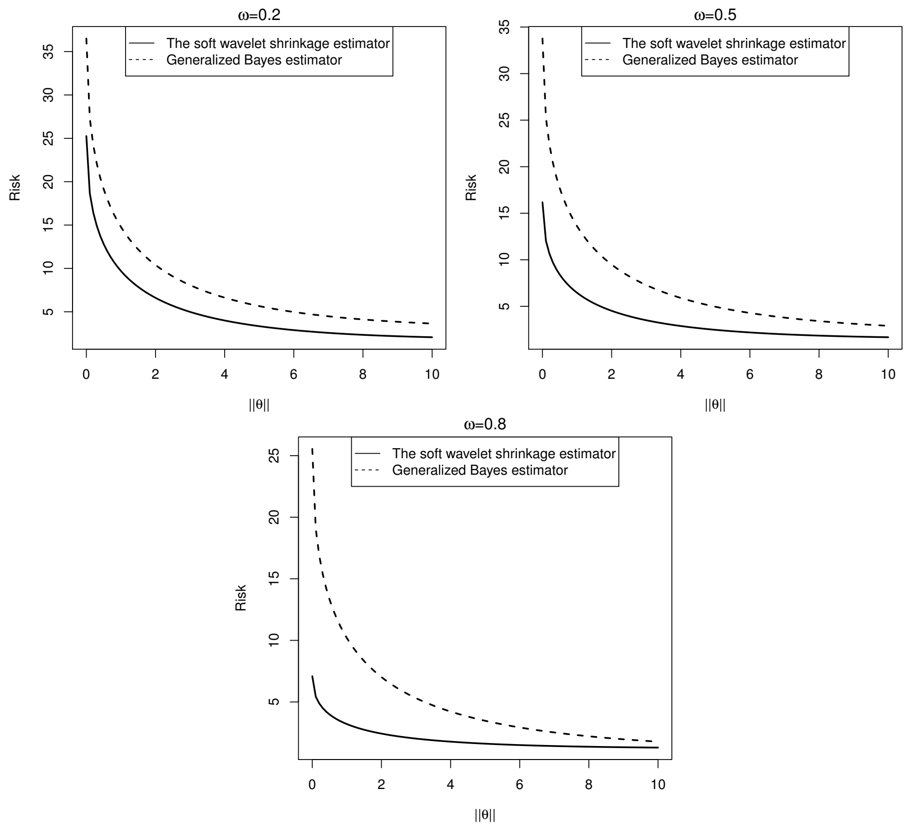

FIGURE 1: Risk plot for the soft wavelet shrinkage estimator and generalized Bayes estimator with $p = 8$ for selected values of ω.

5. 3D Road Network Data Set

In this section, we further investigate the average risk value of the soft wavelet shrinkage estimator for real data set. For this sake, we use the 3D road network data set from Guo et al. (2012). This dataset was constructed by adding elevation information to a 2D road network in North Jutland, Denmark (covering a region of 185×135 km^2). Elevation values where extracted from a publicly available massive Laser Scan Point Cloud for Denmark. This 3D road network was eventually used for benchmarking various fuel and CO2 estimation algorithms. This dataset can be used by any applications that require to know very accurate elevation information of a road network to perform more accurate routing for eco-routing, cyclist routes etc. The dataset contains 4 variables and 434873 observations.

We have implemented a bootstrap analysis to evaluate the risk functions. Table 3 lists the average risk value of the soft wavelet shrinkage estimator for different values of ω. As shown in Table 3, by increasing of the value of ω, the average risk value decreases. In the case of $\omega = 0$, the balanced-LINEX loss function is the basic case of LINEX loss function and it has the most average risk value. Therefore, it can be concluded that the risk values can be reduced by using the balanced-LINEX loss function.

TABLE 3: Average risk value of the soft wavelet shrinkage estimator for 3D road network data set.

δ	$\omega = 0$	$\omega = 0.2$	$\omega = 0.4$	$\omega = 0.6$	$\omega = 0.8$
$\delta^{soft}(X)$	1.610567	1.397633	1.185041	0.9731351	0.7630683

6. Conclusion

In this paper, we consider the generalized Bayes shrinkage estimator of mean vector for multivariate normal distribution under balanced-LINEX loss function. We assume that the random vector X having $N_p(\theta, \Sigma)$ distribution with the unknown covariance matrix Σ. We find minimax and admissible estimator of mean vector based on generalized Bayes estimator. Theoretical findings of this paper are further supported by some numerical analyses. In this regard, the performance evaluation of the proposed class of estimators is checked through a simulation study and real data set.

Acknowledgements

The authors would like to thank the research committee of Persian Gulf University. Also, the authors would like to thank the editors and reviewers for their valuable comments, which greatly improved the readability of this paper.

[Received: December 2020 — Accepted: November 2021]

References

Cao, M. X. & He, D. (2017), 'Admissibility of linear estimators of the common mean parameter in general linear models under a balanced loss function', *Journal of Multivariate Analysis* **153**, 246–254.

Donoho, D. L. & Johnstone, I. M. (1994), 'Ideal spatial adaptation by wavelet shrinkage', *Biometrika* **81**, 425–455.

Fourdrinier, D. & Strawderman, W. E. (2015), 'Robust minimax stein estimation under invariant data-based loss for spherically and elliptically symmetric distributions', *Metrika* **78**(4), 461–484.

Guo, C., Ma, Y., Yang, B., S., J. C. & Kaul, M. (2012), 'EcoMark: Evaluating models of vehicular environmental impact', *SIGSPATIAL/GIS* pp. 269–278.

Huang, S. Y. (2002), 'On a Bayesian aspect for soft wavelet shrinkage estimation under an asymmetric linex loss', *Statistics and Probability Letters* **56**, 171–175.

Jiang, W. & Zhang, C. H. (2009), 'General maximum likelihood empirical Bayes estimation of normal means', *The Annals of Statistics* **37**(4), 1647–1684.

Joly, E. Lugosi, G. & Oliveira, R. I. (2017), 'On the estimation of the mean of a random vector', *Electronic Journal of Statistics* **11**, 440–451.

Jozani, J. M., Leblanc, A. & Marchand, E. (2014), 'On continuous distribution functions, minimax and best invariant estimators, and integrated balanced loss functions', *Canadian Journal of Statistics* **42**, 470–486.

Jozani, M. J., Marchand, É. & Parsian, A. (2006), 'On estimation with weighted balanced-type loss function', *Statistics and Probability Letters* **76**, 733–780.

Jozani, M. J., Marchand, É. & Parsian, A. (2012), 'Bayesian and Robust Bayesian analysis under a general class of balanced loss functions', *Statistical Papers* **53**, 51–60.

Karamikabir, H. & Afshari, M. (2019), 'Wavelet Shrinkage Generalized Bayes Estimation for Elliptical Distribution Parameters under LINEX Loss', *International Journal of Wavelets, Multiresolution and Information Processing* **14**(1), 1950009.

Karamikabir, H. & Afshari, M. (2020), 'Generalized Bayesian Shrinkage and Wavelet Estimation of Location Parameter for Spherical Distribution under Balance-type Loss: Minimaxity and Admissibility,', *Journal of Multivariate Analysis* **177**(1), 104583.

Karamikabir, H. & Afshari, M. (2021), 'New wavelet SURE thresholds of elliptical distributions under the balance loss', *Statistica Sinica* **31**(4), 1829–1852.

Karamikabir, H. Afshari, M. & Arashi, M. (2018), 'Shrinkage estimation of non-negative mean vector with unknown covariance under balance loss', *Journal of Inequalities and Applications* **2018**, 331.

Karamikabir, H., Afshari, M. & Lak, F. (2020), 'Wavelet threshold based on Stein's unbiased risk estimators of restricted location parameter in multivariate normal', *Journal of Applied Statistics* **48**(10), 1712–1729.

Marchand, E. & Strawderman, W. E. (2020), 'On shrinkage estimation for balanced loss functions', *Journal of Multivariate Analysis* **175**, 104558.

Pal, N., Sinha, B. K., Chaudhuri, G. & Chang, C. H. (2007), 'Estimation Of A Multivariate Normal Mean Vector And Local Improvements', *Statistics* **26**(1), 1–17.

Rencher, A. C. & Christensen, W. F. (2012), *Methods of Multivariate Analysis*, third edition edn, John Wiley & Sons.

Rudin, W. (1976), *Principle of Mathematical Analysis*, MacGraw-Hill.

Torehzadeh, S. & Arashi, M. (2014), 'A note on shrinkage wavelet estimation in bayesian analysis', *Statistics and Probability Letters* **84**, 231–234.

Tsukuma, H. & Kubokawa, T. (2015), 'Estimation of the mean vector in a singular multivariate normal distribution', *Journal of Multivariate Analysis* **140**(4), 245–258.

Vidakovic, B. (2009), *Statistical Modelling by Wavelets*, John Wiley and Sons.

Zellner, A. (2009), Bayesian and non-bayesian estimation using balanced loss functions, *in* J. O. Berger & S. S. Gupta, eds, 'Statistical decision theory and related topics V', Springer, New York.

Zinodiny, S., Rezaei, S. & Nadarajah, S. (2017), 'Bayes minimax estimation of the mean matrix of matrix-variate normal distribution under balanced loss function', *Statistics and Probability Letters* **125**, 110–120.