DOI: 10.17516/1997-1397-2020-13-5-608-621
УДК 517.9

Baranchick-type Estimators of a Multivariate Normal Mean Under the General Quadratic Loss Function

Abdenour Hamdaoui*
Department of Mathematics
University of Sciences and Technology, Mohamed Boudiaf, Oran
Laboratory of Statistics and Random Modelisations (LSMA), Tlemcen
Algeria

Abdelkader Benkhaled†
Department of Biology
Mascara University Mustapha Stambouli
Laboratory of Geomatics, Ecology and Environment (LGEO2E)
Mascara, Algeria

Mekki Terbeche‡
Department of Mathematics
University of Sciences and Technology, Mohamed Boudiaf, Oran
Laboratory of Analysis and Application of Radiation (LAAR), USTO-MB
Oran, Algeria

Received 08.04.2020, received in revised form 01.06.2020, accepted 16.07.2020

Abstract. The problem of estimating the mean of a multivariate normal distribution by different types of shrinkage estimators is investigated. We established the minimaxity of Baranchick-type estimators for identity covariance matrix and the matrix associated to the loss function is diagonal. In particular the class of James-Stein estimator is presented. The general situation for both matrices cited above is discussed.

Keywords: covariance matrix, James-Stein estimator, loss function, multivariate gaussian random variable, non-central chi-square distribution, shrinkage estimator.

Citation: A. Hamdaoui, A. Benkhaled, M. Terbeche, Baranchick-type Estimators of a Multivariate Normal Mean Under the General Quadratic Loss Function, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 608–621. DOI: 10.17516/1997-1397-2020-13-5-608-621.

1. Introduction and Preliminaries

The field of estimation of a multivariate normal mean using shrinkage estimators was introduced in [10]. The author showed that the maximum likelihood estimator (MLE) of the mean θ of a multivariate gaussian distribution $N_p\left(\theta, \sigma^2 I_p\right)$ is inadmissible in mean squared sense when the dimension of the parameters space $p \geqslant 3$. In particular, he proved the existence of an estimator which always achieves the smaller total mean squared error regardless of the true θ. Perhaps the best known estimator of such kind is James-Stein's estimator introduced in [7]. This

*abdenour.hamdaoui@yahoo.fr, abdenour.hamdaoui@univ-usto.dz
†benkhaled08@yahoo.fr
‡mekki.terbeche@gmail.com
© Siberian Federal University. All rights reserved

one is a special case of a larger class of estimators known as shrinkage estimators which is a combination of a model with low bias and high variance, and a model with high bias but low variance. In this context we can cite for example Baranchik [2] for his work on the minimaxity of the estimators of the form $\delta_r(X,S) = (1 - r(F)/F)X$ where $F = \|X\|^2/S$, the statistics $S \sim \sigma^2\chi^2_n$ is the estimator of the unknown parameter σ^2 and $r(.)$ is a real mesurable function. Strawderman [12] was interested to study the estimation of the mean vector of a scale mixture of multivariate distribution under squared error loss. He showed the analogous results obtained by Baranchik [2]. Xie et al [13] have introduced a class of semiparametric/parametric shrinkage estimators and established their asymptotic optimality properties. Selahattin et al [9], provided several alternative methods for derivation of the restricted ridge regression estimator (RRRE). The optimal extended balanced loss function (EBLF) estimators and predictors are introduced and derived from [8] and discussed their performances. In [6], the authors considered the model $X \sim N_p\left(\theta, \sigma^2 I_p\right)$ where σ^2 is unknown and estimated by S^2 ($S^2 \sim \sigma^2\chi^2_n$). They studied the following class of shrinkage estimators $\delta_\psi = \delta^{JS} + l(S^2\psi(S^2, \|X\|^2)/\|X\|^2)X$ with l is real parameter. Benkhaled and Hamdaoui [3], have considered the model $X \sim N_p\left(\theta, \sigma^2 I_p\right)$ where σ^2 is unknown. They studied the minimaxity of two different forms of shrinkage estimators of θ: estimators of the form $\delta^\psi = (1 - \psi(S^2, \|X\|^2)S^2/\|X\|^2)X$, and estimators of Lindley-type given by $\delta^\varphi = (1 - \varphi(S^2, T^2)S^2/T^2)(X - \overline{X}) + \overline{X}$.

In this work, we deal with the model $X \sim N_p(\theta, \Sigma)$ and the loss matrix Q where the covariance matrix Σ is known. Our aims is to estimate the unknown parameter θ by shrinkage estimators deduced by the MLE. The paper is organized as follows. In Section 2, we study the standard case $\Sigma = I_p$ and $Q = D = \text{diag}(d_1, d_2, \ldots, d_p)$, we find the explicit formula of the risk function of considered estimators and we treat there minimax property. As a special case, the James-Stein estimator and its risk are also found. In Section 3, we study the considered problem with the generalized matrices Σ and Q. In Section 4, we graphically illustrate risks ratios of the James-Stein estimator and the estimators of Baranchick-type to the MLE for various values of p. We end the manuscript by giving an Appendix which contains technical lemmas used in the proofs of our results.

We recall that if $X \sim N_p\left(\theta, \sigma^2 I_p\right)$, then $\|X\|^2/\sigma^2 \sim \chi^2_p(\lambda)$ where $\chi^2_p(\lambda)$ denotes the non-central chi-square distribution with p degrees of freedom and non-centrality parameter $\lambda = \|\theta\|^2/2\sigma^2$. We also recall the following results that are useful in our proofs.

Definition 1. *For any measurable function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $\chi^2_p(\lambda)$ integrable, we have*

$$E\left[f(\chi^2_p(\lambda))\right] = E_{\chi^2_p(\lambda)}\left[f(U)\right] = \sum_{k=0}^{+\infty}\left[\int_{\mathbb{R}_+} f(u)\chi^2_{p+2k}du\right] P\left(\frac{\lambda}{2}; dk\right),$$

where $P(\lambda/2)$ being the Poisson's distribution of parameter $\lambda/2$ and χ^2_{p+2k} is the central chi-square distribution with $p + 2k$ degrees of freedom.

Lemma 1. *(Stein [11]). Let X be a $N\left(\upsilon, \sigma^2\right)$ real random variable and let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be an indefinite integral of the Lebesgue measurable function, f' essentially the derivative of f. Suppose also that $E\left|f'(X)\right| < +\infty$, Then*

$$E\left[\left(\frac{X - \upsilon}{\sigma^2}\right) f(X)\right] = E\left(f'(X)\right).$$

For the next, if $X \sim N_p(\theta, \Sigma)$, we assume that the loss incurred in estimating θ by δ is the function $L_Q(\delta, \theta) = (\delta - \theta)^t Q(\delta - \theta)$ and the risk function associated to this loss is $R_Q(\delta, \theta) = E_\theta(L_Q(\delta, \theta))$.

2. Results for standard case

Let $X \sim N_p(\theta, I_p)$ be a multivariate gaussian random variable in $\mathbb{R}^p$ and for any estimator δ we take the loss function $L_Q(\delta,\theta) = (\delta-\theta)^t Q(\delta-\theta)$ where $Q = D = \text{diag}\,(d_1, d_2, \dots, d_p)$. It is well known that the MLE of the parameter θ is X and its risk function associated to the loss function L_D is $\sum_{i=1}^{p} d_i = \text{Tr}\,(D)$. Endeed

$$R_D\,(X,\theta) = E\,(L_D\,(X,\theta)) = E\left(\sum_{i=1}^{p} d_i\,(X_i-\theta_i)^2\right) = \sum_{i=1}^{p} d_i E\,(X_i-\theta_i)^2 = \text{Tr}\,(D)\,,$$

because for any i $(i = 1,\dots,p)$, $(X_i-\theta_i)^2 \sim \chi_1^2$ where χ_1^2 is the chi-square distribution with 1 degrees of freedom, then $E_\theta(X_i-\theta_i)^2 = 1$. It is easy to check that the MLE X is minimax, thus any estimator dominates it, is also minimax.

Next, we suppose that $\underline{K} = (K_1,\dots,K_p)$ where K_i $(i = 1,\dots,p)$ are independent Poisson $P\left(\theta_i^2/2\right)$ and $K = \sum_{i=1}^{p} K_i \left(K \sim P\left(\left\|\theta^2\right\|/2\right)\right)$. We give the following Lemma, that can be used in our proofs and its proof is postponed to the Appendix.

Lemma 2. *Let $X \sim N_p\,(\theta, I_p)$ where $X = (X_1,\dots,X_p)^t$ and $\theta = (\theta_1,\dots,\theta_p)^t$. If $p \geqslant 3$, we have*

i) $$E\left(\frac{X_i^2}{\|X\|^2}\right) = E\left(\frac{1+\dfrac{2\theta_i^2}{\|\theta\|^2}K}{p+2K}\right);$$

ii) $$E\left(\frac{X_i^2}{\|X\|^4}\right) = E\left(\frac{1+\dfrac{2\theta_i^2}{\|\theta\|^2}K}{(p-2+2K)\,(p+2K)}\right).$$

2.1. Baranchick-type estimators

In this part, we study the minimaxity of Baranchick-type estimator, which is given by

$$\delta_\psi = \left(1-\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right)X. \tag{1}$$

Proposition 1. *The risk function of the estimator defined in* (1) *under the loss function L_D is*

$$R_D(\delta_\psi,\theta) = \text{Tr}(D) + E\left\{\frac{\sum_{i=1}^{p} d_i\left(1+\dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K}\left[\frac{\psi^2(\chi^2_{p+2K})}{\chi^2_{p+2K}} - 4\psi'(\chi^2_{p+2K}) + 4\frac{\psi(\chi^2_{p+2K})}{\chi^2_{p+2K}}\right]\right\} -$$
$$-\,2\text{Tr}\,(D)\,E\left(\frac{\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right). \tag{2}$$

Proof. We have

$$R_D(\delta_\psi,\theta) = E\left[L_D\left(\delta_\psi,\theta\right)\right] = E\left\{\left(X-\theta-\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)^t D\left(X-\theta-\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} =$$

$$= E\left\{(X-\theta)^t D\,(X-\theta)\right\} + E\left\{\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)^t D\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} -$$

$$-2E\left\{(X-\theta)^t D\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} =$$

$$= \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum_{i=1}^p d_iX_i^2}{\|X\|^4}\right) - 2\sum_{i=1}^p d_iE\left[(X_i-\theta_i)\left[\frac{\psi\left(\|X\|^2\right)X_i}{\|X\|^2}\right]\right].$$

Using Lemma 1, we obtain

$$R_D(\delta_\psi,\theta) = \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum_{i=1}^p d_iX_i^2}{\|X\|^4}\right) - 2\sum_{i=1}^p d_iE\left[\frac{\partial}{\partial X_i}\left[\frac{\psi\left(\|X\|^2\right)X_i}{\|X\|^2}\right]\right] =$$

$$= \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum_{i=1}^p d_iX_i^2}{\|X\|^4}\right) - 4E\left(\frac{\psi'\left(\|X\|^2\right)\sum_{i=1}^p d_iX_i^2}{\|X\|^2}\right) -$$

$$-2\left(\sum_{i=1}^p d_i\right)E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) + 4E\left(\frac{\psi\left(\|X\|^2\right)\sum_{i=1}^p d_iX_i^2}{\|X\|^4}\right) =$$

$$= \mathrm{Tr}\,(D) + E\left\{\frac{\sum_{i=1}^p d_iX_i^2}{\|X\|^2}\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right]\right\} -$$

$$-2\mathrm{Tr}\,(D)\,E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right).$$

From the independence given $\underline{K}$ between $X_i^2/\|X\|^2$ and $\|X\|^2$ for $i=1,\ldots,p$, we get

$$R_D(\delta_\psi,\theta) = \mathrm{Tr}\,(D) +$$

$$+E\left\{\sum_{i=1}^p d_iE\left(\frac{X_i^2}{\|X\|^2}|\underline{K}\right)E\left[\left(\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right)|\underline{K}\right]\right\} -$$

$$-2\left(\sum_{i=1}^p d_i\right)E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right).$$

Using the Lemma 2, we have

$$
\begin{aligned}
R_D(\delta_\psi,\theta) = \operatorname{Tr}(D) + \\
&+ E\left\{\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K} E\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}|\underline{K}\right]\right\} - \\
&- 2\operatorname{Tr}(D)\, E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) = \\
&= \operatorname{Tr}(D) - 2\operatorname{Tr}(D)\, E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) + \\
&+ E\left\{E\left\{\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K}\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}|\underline{K}\right]\right\}\right\}.
\end{aligned}
$$

From Definition 1 and using properties of conditional expectation we have, for any two measurable functions G and H, $E\left[G\left(\|X\|^2\right)\right] = E\left[G\left(\chi^2_{p+2K}\right)\right]$ and $E\left\{E\left[H(K)\,|\underline{K}\right]\right\} = E\left[H(K)\right]$, where $K \sim P\left(\|\theta\|^2/2\right)$, thus we get the desired result. □

Note that the classical result of minimaxity of Baranchick-type estimators which is obtained for the loss function $L(\delta,\theta) = \sum_{i=1}^{p}(\delta_i-\theta_i)^2$ (i.e. $d_i = 1$ for any $i = 1,\ldots,p$), is also available and it is established in the following Theorem.

Theorem 1. *Assume that δ_ψ is given in (1) with $p \geqslant 3$. Under the loss function L_D with $\frac{\operatorname{Tr}(D)}{\max_{1\leqslant i\leqslant p}(d_i)} \geqslant 2$, if*

i) $\psi(.)$ is monotone non-decreasing function;

ii) $0 \leqslant \psi(.) \leqslant 2\left(\frac{\operatorname{Tr}(D)}{\max_{1\leqslant i\leqslant p}(d_i)} - 2\right)$,

then δ_ψ is minimax.

Proof. From formula (2), we have

$$
\begin{aligned}
R_D(\delta_\psi,\theta) \leqslant \operatorname{Tr}(D) + \\
&+ E\left\{\frac{-4\psi'\left(\chi^2_{p+2K}\right)\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K} + \max_{1\leqslant i\leqslant p}(d_i)\left[\frac{\psi^2\left(\chi^2_{p+2K}\right)+4\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right]\right\} - \\
&- 2\operatorname{Tr}(D)\, E\left(\frac{\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right) \leqslant
\end{aligned}
$$

$$\leqslant \mathrm{Tr}\,(D) + E\left[\frac{-4\psi'\left(\chi^2_{p+2K}\right)\sum_{i=1}^{p} d_i\left(1+\dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K}\right] +$$

$$+ E\left\{\frac{\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\left[\max_{1\leqslant i\leqslant p}(d_i)\left[\psi\left(\chi^2_{p+2K}\right)+4\right] - 2\mathrm{Tr}\,(D)\right]\right\}.$$

Then, a sufficient condition for that δ_ψ is minimax is that $\psi\,(.)$ is a positive monotone non-decreasing function and $\max_{1\leqslant i\leqslant p}(d_i)\left[\psi\left(\chi^2_{p+2K}\right)+4\right] - 2\mathrm{Tr}\,(D) \leqslant 0$. Which are equivalent to

$$0 \leqslant \psi\,(.) \leqslant 2\left(\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)} - 2\right)$$

and $\psi\,(.)$ is monotone non-decreasing. □

Example 1. Let the shrinkage functions $\psi^{(1)}\left(\|X\|^2\right) = \|X\|^2/\left(\|X\|^2+1\right)$, $\psi^{(2)}\left(\|X\|^2\right) = = 1 - exp(-\|X\|^2)$ and the matrices $D^{(1)} = \mathrm{diag}\,(d_1 = 1, d_2 = 1/2, \ldots, d_p = 1/p)$ with $p \geqslant 7$ and $D^{(2)} = \mathrm{diag}\,(d_1 = 1/2, d_2 = 2/3, \ldots, d_p = p/p+1)$ with $p \geqslant 4$. It is clear that the functions $\psi^{(1)}\,(.)$ and $\psi^{(2)}\,(.)$ satisfie conditions of Theorem 1. Then the estimators $\delta_{\psi^{(1)}}$ and $\delta_{\psi^{(2)}}$ are minimax for $p \geqslant 7$ under the loss function $L_D^{(1)}$ and are minimax for $p \geqslant 4$ under the loss function $L_D^{(2)}$.

Now, we discuss the special case where $\psi\,(.) = a$ with a is a positive constant.

2.2. James-Stein estimator

Consider the estimator $\delta_a = \left(1 - a/\|X\|^2\right)X = X - \left(a/\|X\|^2\right)X$, where a is a real parameter that can depend on p. Using the Proposition 1, the risk function of the estimator δ_a is

$$R_D(\delta_a,\theta) = \mathrm{Tr}\,(D) + a\,(a+4)\,E\left(\frac{\sum_{i=1}^{p} d_i\left(1+\dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)\,(p+2K)}\right) - 2a\mathrm{Tr}\,(D)\,E\left(\frac{1}{p-2+2K}\right). \quad (3)$$

Proposition 2. *Under the loss function L_D with $p \geqslant 3$ and $\dfrac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)} \geqslant 2$, we have*

i) a sufficient condition for that δ_a dominates the MLE X is

$$0 \leqslant a \leqslant 2\left(\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)} - 2\right);$$

ii) the optimal value of a that minimizes the risk function $R_D(\delta_a,\theta)$ is

$$\widehat{a} = \frac{\mathrm{Tr}\,(D)\,E\left(\dfrac{1}{(p-2+2K)}\right)}{\alpha} - 2,$$

where $\alpha = E\left(\sum_{i=1}^{p} d_i\left(1+\left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)(p+2K)\right)$.

Proof. i) From formula (3), a sufficient condition so that δ_a dominating the MLE X is

$$a(a+4)E\left(\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)(p+2K)}\right)-2a\mathrm{Tr}(D)E\left(\frac{1}{p-2+2K}\right)\leqslant 0.$$

As

$$E\left(\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)(p+2K)}\right)\leqslant \max_{1\leqslant i\leqslant p}(d_i)\,E\left(\frac{1}{p-2+2K}\right),$$

thus, a sufficient condition so that δ_a dominates the MLE X is

$$a\left[(a+4)\max_{1\leqslant i\leqslant p}(d_i)-2\mathrm{Tr}(D)\right]E\left(\frac{1}{p-2+2K}\right)\leqslant 0,$$

which is equivalent to the desired result.

ii) Using the convexity of the risk function $R_D(\delta_a,\theta)$ on a, one can easily show that the optimal value of a that minimizes the risk function $R_D(\delta_a,\theta)$ is $\widehat{a}=(\mathrm{Tr}(D)E(1/(p-2+2K))/\alpha)-2$, where $\alpha = E\left(\sum_{i=1}^{p} d_i\left(1+\left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)(p+2K)\right)$. □

For $a=\widehat{a}$ we obtain the James-Stein estimator $\delta_{JS}\left(=\delta_{\widehat{a}}=\left(1-\widehat{a}/\|X\|^2\right)X\right)$ which minimizes the risk function of estimators δ_a, so that from formula (3), the risk function of the James-Stein estimator δ_{JS} under the loss function L_D is

$$R_D(\delta_{JS},\theta)=\mathrm{Tr}(D)-\frac{\left[\mathrm{Tr}(D)E\left(\frac{1}{p-2+2K}\right)-2\alpha\right]^2}{\alpha}, \tag{4}$$

As the constant α is non-negative and using the formula (4), it is clear that the James-Stein estimator δ_{JS}, has a risk less than $\mathrm{Tr}(D)$, then δ_{JS} is minimax.

3. The case of generalized Σ and Q

Let $X\sim N_p(\theta,\Sigma)$ and the loss function $L_Q(\delta,\theta)=(\delta-\theta)^tQ(\delta-\theta)$ where the covariance matrix Σ is known and $\Sigma^{1/2}Q\Sigma^{1/2}$ is diagonalizable matrix. Take the change of variables $Y=P\Sigma^{-1/2}X$ where P is an orthogonal matrix ($PP^t=I_p$) that diagonalizes the matrix $\Sigma^{1/2}Q\Sigma^{1/2}$ such as $P\Sigma^{1/2}Q\Sigma^{1/2}P^t=D^*=\mathrm{diag}(a_1,\ldots,a_p)$. Then we have $Y\sim N_p(\nu,I_p)$ with $\nu=P\Sigma^{-1/2}\theta$. Thus the risk function of the MLE X associated to the loss function L_Q is $\sum_{i=1}^{p} a_i=\mathrm{Tr}(D^*)$. Endeed

$$\begin{aligned}R_Q(X,\theta)&=E\left[(X-\theta)^tQ(X-\theta)\right]=E\left\{\left[\Sigma^{1/2}P^{-1}(Y-\nu)\right]^tQ\left[\Sigma^{1/2}P^{-1}(Y-\nu)\right]\right\}=\\&=E\left\{(Y-\nu)^tP\Sigma^{1/2}Q\Sigma^{1/2}P^t(Y-\nu)\right\}=E\left\{(Y-\nu)^tD^*(Y-\nu)\right\}=\\&=\sum_{i=1}^{p}a_iE\left[(Y_i-\nu_i)^2\right]=\mathrm{Tr}(D^*),\end{aligned}$$

because for any i $(i = 1, \dots, p)$ $(Y_i - \nu_i)^2 \sim \chi_1^2$ where χ_1^2 is the chi-square distribution with 1 degrees of freedom, thus $E\left(Y_i - \nu_i\right)^2 = 1$. As the MLE X is minimax, then any estimator dominates it, is also minimax.

3.1. Baranchik-type estimators

Now, consider the estimator given by

$$\delta_\phi = \left(1 - \frac{\phi\left(X^t \Sigma^{-1} X\right)}{X^t \Sigma^{-1} X}\right) X. \tag{5}$$

Proposition 3. *Under the loss function L_Q the risk function of the estimator $\delta_\phi$is*

$$R_Q(\delta_\phi, \theta) = \operatorname{Tr}\left(D^*\right) + E\left\{\frac{\sum\limits_{i=1}^{p} a_i \left(1 + \frac{2\theta_i^2}{\|\theta\|^2} K\right)}{p + 2K}\left[\frac{\phi^2\left(\chi_{p+2K}^2\right)}{\chi_{p+2K}^2} - 4\phi'\left(\chi_{p+2K}^2\right) + 4\frac{\phi\left(\chi_{p+2K}^2\right)}{\chi_{p+2K}^2}\right]\right\} -$$
$$- 2\operatorname{Tr}\left(D^*\right) E\left(\frac{\phi\left(\chi_{p+2K}^2\right)}{\chi_{p+2K}^2}\right),$$

where $K \sim P\left(\|\nu\|^2/2\right)$.

Proof.

$$R_Q(\delta_\phi, \theta) = E\left\{\left[\left(1 - \frac{\phi\left(X^t \Sigma^{-1} X\right)}{X^t \Sigma^{-1} X}\right) X - \theta\right]^t Q \left[\left(1 - \frac{\phi\left(X^t \Sigma^{-1} X\right)}{X^t \Sigma^{-1} X}\right) X - \theta\right]\right\} =$$
$$= E\left\{\left[(X - \theta) - \frac{\phi\left(X^t \Sigma^{-1} X\right)}{X^t \Sigma^{-1} X} X\right]^t Q \left[(X - \theta) - \frac{\phi\left(X^t \Sigma^{-1} X\right)}{X^t \Sigma^{-1} X} X\right]\right\}.$$

Using the change variable $Y = P\Sigma^{-1/2}X$ where P is an orthogonal matrix and P diagonalizes the matrix $\Sigma^{1/2} Q \Sigma^{1/2}$, then $Y \sim N_p\left(\nu, I_p\right)$ with $\nu = P\Sigma^{-1/2}\theta$ and

$$R_Q(\delta_\phi, \theta) = E\left\{\left[\Sigma^{1/2} P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^t Y\right)}{Y^t Y} Y\right]\right]^t Q \left[\Sigma^{1/2} P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^t Y\right)}{Y^t Y} Y\right]\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(Y^t Y\right)}{Y^t Y} Y\right]^t \left(P^{-1}\right)^t \Sigma^{1/2} Q \Sigma^{1/2} P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^t Y\right)}{Y^t Y} Y\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2} Y\right]^t P\Sigma^{1/2} Q \Sigma^{1/2} P^t \left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2} Y\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2} Y\right]^t D^* \left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2} Y\right]\right\} = R_{D^*}(\delta_\phi^*, \theta),$$

where $\|.\|$ is the usual euclidean norm in $\mathbb{R}^p$, $P\Sigma^{1/2} Q \Sigma^{1/2} P^t = D^* = \operatorname{diag}\left(a_1, \dots, a_p\right)$ and $\delta_\phi^* = \left(1 - (\phi(\|Y\|^2/\|Y\|^2)\right) Y$. From Proposition 1, we obtain the desired result. □

Theorem 2. *Assume that δ_ϕ is given by (5) where $p \geqslant 3$. Under the loss function L_Q with $\frac{\mathrm{Tr}(D^*)}{\max_{1\leqslant i\leqslant p}(a_i)} \geqslant 2$, if*

i) $\phi(.)$ is monotone non-decreasing;

ii) $0 \leqslant \phi(.) \leqslant 2\left(\frac{\mathrm{Tr}(D^*)}{\max_{1\leqslant i\leqslant p}(a_i)} - 2\right)$,

then δ_ϕ is minimax.

The proof is the same given for the Theorem 1.

3.2. James-Stein estimator

Consider the estimator $\delta_b = \left(1 - b/\left(X^t\Sigma^{-1}X\right)\right)X$. Using the Proposition 3, one can show easily that the risk function of the estimator δ_b under the loss function L_Q is.

$$R_Q(\delta_b,\theta) = \mathrm{Tr}(D^*) + b(b+4)E\left(\frac{\sum_{i=1}^{p} a_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)(p+2K)}\right) - 2b\mathrm{Tr}(D^*)E\left(\frac{1}{(p-2+2K)}\right),$$

where $K \sim P\left(\|\nu\|^2/2\right)$. From the last formula, we deduce immediately that, a sufficient condition for that δ_b dominating the MLE X is $0 \leqslant b \leqslant 2\left((\mathrm{Tr}(D^*)/\max_{1\leqslant i\leqslant p} a_i) - 2\right)$, and the optimal value of b that minimizes the risk function $R_Q(\delta_b,\theta)$ is

$$\widehat{b} = \frac{\mathrm{Tr}(D^*)E\left(\frac{1}{(p-2+2K)}\right)}{\beta} - 2,$$

where $\beta = E\left(\sum_{i=1}^{p} a_i\left(1+\left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)(p+2K)\right)$.

For $b = \widehat{b}$ we obtain the James-Stein estimator $\delta^*_{JS} = \delta_{\widehat{b}} = \left(1 - \widehat{b}/\left(X^t\Sigma^{-1}X\right)\right)X$ which minimizes the risk function of δ_b. Its risk function associated to the loss function L_Q is

$$R_Q(\delta^*_{JS},\theta) = \mathrm{Tr}(D^*) - \frac{\left[\mathrm{Tr}(D^*)E\left(\frac{1}{p-2+2K}\right) - 2\beta\right]^2}{\beta}. \tag{6}$$

From formula (6), we note that δ^*_{JS} dominates the MLE X, thus δ^*_{JS} is minimax.

4. The simulation results

In this section we take the model $X \sim N_p(\theta, I_p)$ where $\theta = (\theta_1, \theta_1, \ldots, \theta_1)^t$ and we recall the estimators of Baranchick-type and the matrices $D^{(1)}$ and $D^{(2)}$ given in Example 1, i.e., $\delta_{\psi^{(1)}} = \left(1 - \psi^{(1)}\left(\|X\|^2\right)/\|X\|^2\right)X$, $\delta_{\psi^{(2)}} = \left(1 - \psi^{(2)}\left(\|X\|^2\right)/\|X\|^2\right)X$ with $\psi^{(1)}\left(\|X\|^2\right) = \|X\|^2/\left(\|X\|^2+1\right)$, $\psi^{(2)}\left(\|X\|^2\right) = 1 - exp\left(-\|X\|^2\right)$, $D^{(1)} = \mathrm{diag}\,(d_1 = 1, d_2 = 1/2, \ldots, d_p = 1/p)$

and $D^{(2)} = \mathrm{diag}\,(d_1 = 1/2, d_2 = 2/3, \ldots, d_p = p/p+1)$. We also recall the form of the James-Stein estimator $\delta_{JS}\left(= \delta_{\widehat{a}} = \left(1 - \widehat{a}\,/\,\|X\|^2\right) X\right)$, where $\widehat{a} = (\mathrm{Tr}\,(D)\, E\,(1/\,(p-2+2K))/\alpha) - 2$ and $\alpha = E\left(\sum_{i=1}^{p} d_i \left(1 + \left(2\theta_i^2/\|\theta\|^2\right) K\right)/(p-2+2K)\,(p+2K)\right)$. We graph the risks ratios of estimators cited above, to the MLE associated the the losses functions $L_{D^{(1)}}$ and $L_{D^{(2)}}$ denoted respectively: $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ as function of $\lambda = \theta_1^2$ for various values of p.

In Figs. 1–4, we note that the risks ratios $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ are less than 1, thus the estimators δ_{JS}, $\delta_{\psi^{(1)}}$ and $\delta_{\psi^{(2)}}$ are minimax for $p = 8$ and $p = 12$ under the loss function $L_{D^{(1)}}$, and also minimax for $p = 4$ and $p = 6$ under the loss function $L_{D^{(2)}}$.

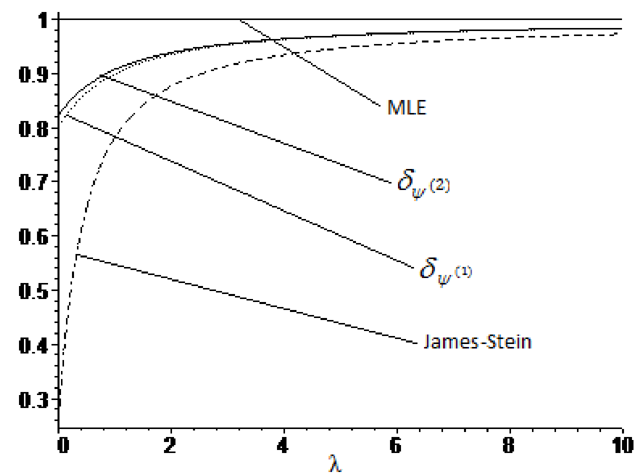

Fig. 1. Graph of risks ratios $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ as function of $\lambda = \theta_1^2$ for $p = 8$ under the loss function $L_{D^{(1)}}$

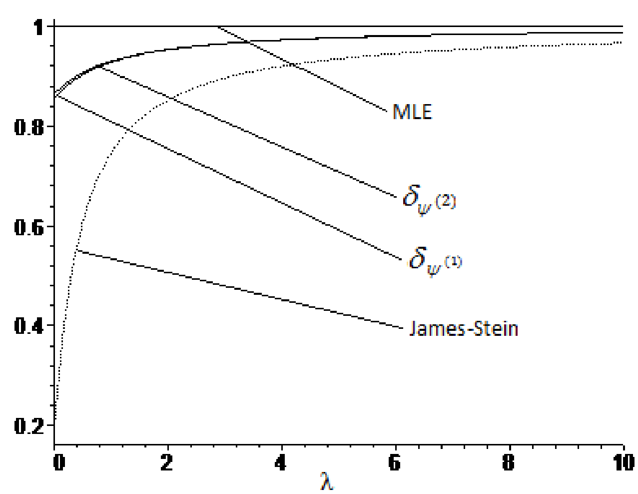

Fig. 2. Graph of risks ratios $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ as function of $\lambda = \theta_1^2$ for $p = 12$ under the loss function $L_{D^{(1)}}$

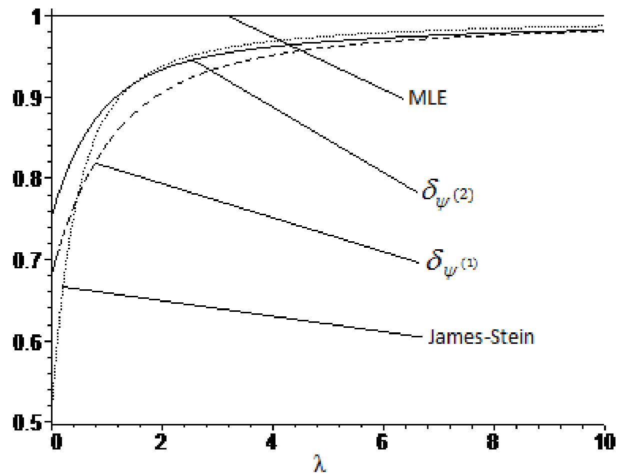

Fig. 3. Graph of risks ratios $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ as function of $\lambda=\theta_1^2$ for $p=4$ under the loss function $L_{D^{(2)}}$

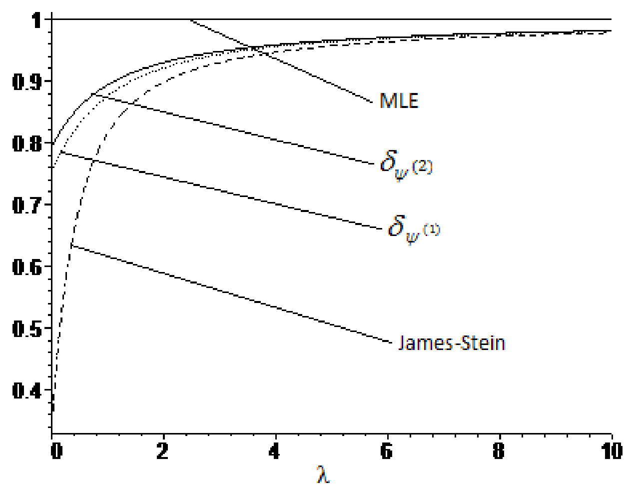

Fig. 4. Graph of risks ratios $R(\delta_{JS},\theta)/R(X,\theta)$, $R(\delta_{\psi^{(1)}},\theta)/R(X,\theta)$ and $R(\delta_{\psi^{(2)}},\theta)/R(X,\theta)$ as function of $\lambda=\theta_1^2$ for $p=6$ under the loss function $L_{D^{(2)}}$

5. Appendix

Lemma 3 (Bock [5]). *Let $X\sim N_p(\theta,I_p)$ where $X=(X_1,\ldots,X_p)^t$ and $\theta=(\theta_1,\ldots,\theta_p)^t$, then, For any measurable function $h:[0,+\infty[\rightarrow\mathbb{R}$, we have*

$$E\left(h\left(\|X\|^2\right)X_i^2\right)=E\left[h\left(\chi^2_{p+2}\left(\|\theta\|^2\right)\right)\right]+\theta_i^2E\left[h\left(\chi^2_{p+4}\left(\|\theta\|^2\right)\right)\right].$$

where $K\sim P\left(\|\theta\|^2/2\sigma^2\right)$ being the Poisson's distribution of parameter $\|\theta\|^2/2\sigma^2$.

Lemma 4 (Bock [5]). *Let f be a real-valued mesurable function defined on the integer. Let $K \sim P(\lambda/2)$ being the Poison's distribution of parameter $\lambda/2$. Then*

$$\lambda E\left[f(K)\right] = E\left[2Kf(K-1)\right],$$

if both sides exist.

Proof Lemma 2. i) Using Lemma 3 and the Definition 1, we obtain

$$E\left(\frac{X_i^2}{\|X\|^2}\right) = E_{\chi^2_{p+2}(\|\theta\|^2)}\left(\frac{1}{u}\right) + \theta_i^2 E_{\chi^2_{p+4}(\|\theta\|^2)}\left(\frac{1}{u}\right) = E\left(\frac{1}{p+2K}\right) + \theta_i^2 E\left(\frac{1}{p+2+2K}\right),$$

where $K \sim P\left(\|\theta\|^2/2\right)$ being the Poisson's distribution of parameter $\|\theta\|^2/2$.
From Lemma 4, we have

$$E\left(\frac{X_i^2}{\|X\|^2}\right) = E\left(\frac{1}{p+2K}\right) + \frac{\theta_i^2}{\|\theta\|^2}E\left(\frac{2K}{p+2K}\right) = E\left(\frac{1+2\frac{\theta_i^2}{\|\theta\|^2}K}{p+2K}\right).$$

ii) Using Lemma 3 and the Definition 1, we obtain

$$E\left(\frac{X_i^2}{\|X\|^4}\right) = E_{\chi^2_{p+2}(\|\theta\|^2)}\left(\frac{1}{u^2}\right) + \theta_i^2 E_{\chi^2_{p+4}(\|\theta\|^2)}\left(\frac{1}{u^2}\right) =$$
$$= E\left(\frac{1}{(p-2+2K)(p+2K)}\right) + \theta_i^2 E\left(\frac{1}{(p+2K)(p+2+2K)}\right),$$

where $K \sim P\left(\|\theta\|^2/2\right)$. From Lemma 4, we have

$$E\left(\frac{X_i^2}{\|X\|^2}\right) = E\left(\frac{1}{(p-2+2K)(p+2K)}\right) + \frac{\theta_i^2}{\|\theta\|^2}E\left(\frac{2K}{(p-2+2K)(p+2K)}\right) =$$
$$= E\left(\frac{1+2\frac{\theta_i^2}{\|\theta\|^2}K}{(p-2+2K)(p+2K)}\right).$$

□

Conclusion

Stein [10], has started to study the estimation of the mean θ of a multivariate gaussian random $N_p\left(\theta, \sigma^2 I_p\right)$ in $\mathbb{R}^p$, by the shrinkage estimators deduced from the usual estimator. Many authors continued to work in this field. The majority among them have studied the minimaxity of these estimators under the usual quadratic risk function, we cite for example [5,7]. Other authors research the stability of the minimaxity property in the case where the dimension of the parameter space and the sample size are large, we refer to [3,6]. In this work we studied the minimaxity of Baranchick-type estimators, relatively to the general loss function. We showed similar results to those found in the classical case. An idea would be to see whether one can obtain similar results of the minimaxity and the asymptotic behavior of risk ratios in the general case of the symmetrical spherical models.

The authors are extremely grateful to the editor and the referees for carefully reading the paper and for valuable comments and helpful remarks and suggestions which greatly improved the presentation of this paper. This research is supported by the Thematic Research Agency in Science and Technology (ATRST-Algeria).

References

[1] A.J.Baranchick, Multiple Regression and estimation of the mean of a multivariate normal distribution, Stanford Univ., Technical Report, no. 51, 1964.

[2] A.J.Baranchick, A family of minimax estimators of the mean of a multivariate normal distribution, *The Annals of Mathematical Statistics*, **41**(1970), no. 2, 642–645.

[3] A.Benkhaled, A.Hamdaoui, General classes of shrinkage estimators for the multivariate normal mean with unknown variance: minimaxity and limit of risks ratios, *Kragujevac J. Math.*, **46**(2019), no. 2, 193–213.

[4] D.Benmansour, A.Hamdaoui, Limit of the ratio of risks of James-Stein estimators with unknown variance, *Far East J. Theo. Stat.*, **36**(2011), no. 1, 31–53.

[5] M.E.Bock, Minimax estimators of the mean of a multivariate normal distribution, *Ann. Statist.*, **3**(1975), no. 1, 209–218.

[6] A.Hamdaoui, D.Benmansour, Asymptotic properties of risks ratios of shrinkage estimators, *Hacet. J. Math. Stat.*, **44**(2015), no. 5, 1181–1195.

[7] W.James, C.Stein, Estimation with Quadratic Loss, Proc. 4th Berkeley Symp, Math. Statist. Prob, Univ of California Press, Berkeley, Vol. 1, 1961, 361–379.

[8] K.Selahattin, D.Issam, The optimal extended balanced loss function estimators, *J. Comput. Appl. Math.*, **345**(2019), 86–98.

[9] K.Selahattin, S.Sadullah, M.RÖzkale, H.Güler, More on the restricted ridge regression estimation, *J. Stat. Comput. Simul.*, **81**(2011), 1433–1448.

[10] C.Stein, Inadmissibilty of the usual estimator for the mean of a multivariate normal distribution, Proc. 3th Berkeley Symp, Math. Statist. Prob. Univ. of California Press, Berkeley, Vol. 1, 1956, 197–206.

[11] C.Stein, Estimation of the mean of a multivariate normal distribution, *Ann. Statis.*, **9**(1981), no. 6, 1135–1151.

[12] W.E.Strawderman, Minimax estimation of location parameters for certain spherically symmetric distribution, *J. Multivariate Anal.*, **4**(1974), no. 1, 255–264.

[13] X.Xie, S.C.Kou, L.Brown, Optimal shrinkage estimators of mean parameters in family of distribution with quadratic variance, *Ann. Statis.*, **44**(2016), no. 2, 564–597.

Об оценках решений задачи расщепления для некоторых многомерных дифференциальных уравнений в частных производных

Абденур Хамдауи
Университет науки и технологии Оран Мохамед-Будиаф
Оран, Алжир
Университет Тлемсена (LSMA)
Тлемсен, Алжир
Абделькадер Бенхалед
Университет Мюстафа Стамбули (LGEO2E)
Маскара, Алжир
Мекки Тербече
Университет науки и технологии Оран Мохамед-Будиаф (LAAR), USTO-MB
Оран, Алжир

Аннотация. Исследована проблема оценки среднего многомерного нормального распределения различными типами оценок усадки. Мы установили минимаксность оценок типа Баранчика для единичной ковариационной матрицы, а матрица, связанная с функцией потерь, является диагональной. В частности, представлен класс оценки Джеймса-Стейна. Обсуждается общая ситуация для обеих упомянутых выше матриц.

Ключевые слова: ковариационная матрица, оценка Джеймса-Стейна, функция потерь, многомерная гауссовская случайная величина, нецентральное распределение хи-квадрат, оценка усадки.