Metrika (2010) 72:51–57
DOI 10.1007/s00184-009-0240-9

Estimation of the location parameter under LINEX loss function: multivariate case

M. Arashi · S. M. M. Tabatabaey

Received: 8 June 2006 / Published online: 26 February 2009

Abstract The Baysian estimation of the mean vector θ of a p-variate normal distribution under linear exponential (LINEX) loss function is studied when as a special restricted model, it is suspected that for a $p \times r$ known matrix Z the hypothesis $\theta = Z\beta$, $\beta \in \Re^r$ may hold. In this area we show that the Bayes and empirical Bayes estimators dominate the unrestricted estimator (when nothing is known about the mean vector θ).

Keywords Moment generating function · Empirical Bayes · LINEX loss function · Restricted model

1 Introduction

A symmetric loss function assumes that positive and negative errors are equally serious. In situations where negative bias and positive bias of the same magnitude have different importance, symmetric loss function seems inappropriate.

As an asymmetric loss function, consider the following linear exponential (LINEX) loss function with the scale parameter b and the p-vector shape parameter a

$$L(\delta(X), \theta) = b\left\{e^{a'[\delta(X)-\theta]} - a'[\delta(X)-\theta] - 1\right\}, \tag{1.1}$$

where $\delta(X)$ is an estimator of p-vector parameter θ based on the random vector X.

M. Arashi (✉)
Faculty of Mathematics, Shahrood University of Technology,
P.O. Box 316, 3619995161 Shahrood, Iran
e-mail: m_arashi_stat@yahoo.com

S. M. M. Tabatabaey
Department of Statistics, School of Mathematical Sciences,
Ferdowsi University of Mashhad, Mashhad, Iran

Let X be a random vector from $N_p(\theta, \Sigma)$ where the mean vector θ is unknown and the positive definite covariance matrix Σ is known. When nothing is known about the mean vector θ, then the maximum likelihood estimator (MLE)of θ is given by X; which is an unbiased estimator of the mean vector θ. It is well-known that James and Stein (1961) considered a decision—theoretic approach to the estimation of θ under squared error loss (SEL) when Σ is known and θ may be equal to vector zero (0), while Efron and Morris (1972, 1976) considered the same problem when Σ is unknown and θ may be equal to 0 using an empirical Bayes (EB) approach. Also Srivastava and Saleh (2005) have studied the empirical Bayes estimator (EBE) of θ under SEL where the hypothesis $H_0 : \theta = Z\beta$ may hold.

In this paper, we consider the estimation of θ under the LINEX loss function when θ may belong to a sub-space, $\theta = Z\beta$, where Z is p×r matrix of known constants with rank r and $\beta \in \Re^r$. The estimator in this situation is called restricted estimator. For a complete discussion of restricted estimators in different classical models, see Saleh (2006) and the references therein.

2 Model modification

Consider the restricted model in which for the known p×r matrix Z of rank r and $\beta \in \Re^r$, the hypothesis $H_0 : \theta = Z\beta$ holds, then the MLE of θ is given by

$$\begin{aligned}\hat{\theta} &= Z\hat{\beta} \\ &= Z(Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-1}X. \end{aligned} \tag{2.1}$$

Now suppose the prior distribution of θ is given by

$$\theta \mid \beta \sim N_p(Z\beta, \alpha\Sigma), \tag{2.2}$$

for a known positive parameter α.

Given θ, $X \sim N_p(\theta, \Sigma)$, the marginal distribution of X is $N_p(Z\beta, (1+\alpha)\Sigma)$ and the conditional distribution of θ given X is

$$\theta \mid X \sim N_p\left(\mu, \frac{\alpha}{1+\alpha}\Sigma\right). \tag{2.3}$$

where $\mu = \frac{1}{1+\alpha}Z\beta + \frac{\alpha}{1+\alpha}X$.

The Bayes estimator under LINEX loss function is that one which minimizes the posterior expectation of the loss function in (1.1). Using (2.2) we have

$$\begin{aligned}\frac{\partial E_{\theta|X}L[\delta(X)-\theta]}{\partial\delta(X)} &= b\left\{\frac{\partial e^{a'\delta(X)}}{\partial\delta(X)}M_{\theta|X}(-a) - \frac{\partial a'\delta(X)}{\partial\delta(X)}\right\} \\ &= b\left\{a\exp\left[a'\delta(X) - a'\mu + \frac{\alpha}{2(1+\alpha)}\,a'\Sigma a\right] - a\right\}, \end{aligned} \tag{2.4}$$

where $M_Y(.)$ is the moment generating function of Y.

Then for any p-vector valued a, the Bayes estimator for θ in the restricted model is $\hat{\delta}_B(X)$ where

$$a'\hat{\delta}_B(X) = a'\mu - \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a. \tag{2.5}$$

As an example, for $a = (0, \ldots, 1, 0, \ldots, 0)$ (ith element is 1), the ith element of $\hat{\delta}_B(X)$ is

$$\hat{\delta}_{B_i}(X) = \mu_i - \frac{\alpha}{2(1+\alpha)}\, \sigma_{ii},$$

where μ_i is the ith element of μ.

Under the restricted model in which the hypothesis $H_0 : \theta = Z\beta$ holds (θ has a distribution as $\pi(\theta \mid \beta)$, when the hyperparameter (β) from a conjugate prior $\pi(\theta \mid \beta)$ are unavailable, we use an estimate of the hyperparameter such as the MLE; that is an empirical Bayes problem.

In order to obtain the EBE in the restricted model for θ, all that has to be done is to use $\hat{\mu}$, replacing $Z\beta$ by the MLE of θ, (2.1). Therefore, using (2.4) for any p-vector a, the EBE is $\hat{\delta}_{EB}(X)$ where

$$\begin{aligned} a'\hat{\delta}_{EB}(X) &= a'\hat{\mu} - \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a \\ &= \frac{1}{1+\alpha}\, a'\left[Z(Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-1} + \alpha I_p\right]X \\ &\quad - \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a. \end{aligned} \tag{2.6}$$

Now consider the special cases in which

(1) Let $r = p$ and Z is positive definite, we obtain the EBE, $\hat{\delta}_{EB1}(X)$ for θ as

$$a'\hat{\delta}_{EB1}(X) = a'X - \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a. \tag{2.7}$$

(2) Let $Z = (I_r, 0)'$; when all the last $(p - r)$ components of θ are zero. For the situation where

$$D_{p\times p} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

we have $\hat{\theta} = DX$. Therefore the EBE for θ is

$$a'\hat{\delta}_{EB2}(X) = \frac{1}{1+\alpha} a'\left[D + \alpha I_p\right]X - \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a. \tag{2.8}$$

3 Bayes risk

In this section we derive the Bayes risks of the estimators in the previous section.

The Bayes risk of the unrestricted estimator (UE) X under the LINEX loss function is

$$R(X,\theta) = b\left\{\exp\left[\frac{1}{2}\, a'\Sigma a\right] - 1\right\}. \tag{3.1}$$

Using (2.1), (2.2) and (2.5) the Bayes risk of $\hat{\delta}_B(X)$ is given by

$$\begin{aligned} R(\hat{\delta}_B(X),\theta) &= E_\theta\left\{E_{X|\theta} L[\hat{\delta}_B(X)-\theta]\right\} \\ &= b\left[\frac{\alpha}{2(1+\alpha)}\, a'\Sigma a - 1\right]. \end{aligned} \tag{3.2}$$

Taking $Y = Z(Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-1} + \alpha I_p$, we obtain

$$Y\Sigma Y' = (2\alpha+1)Z(Z'\Sigma^{-1}Z)^{-1}Z' + \alpha^2\Sigma, \tag{3.3}$$

Using (2.6) and (3.3) the Bayes risk of the EBE is given by

$$\begin{aligned} R(\hat{\delta}_{EB}(X),\theta) &= b\left(\exp\left\{\frac{1}{2(1+\alpha)}\, a'[Y\Sigma Y' - \alpha\Sigma - 2\alpha Y\Sigma]a\right\}\right. \\ &\quad \left. + \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a - 1\right) \\ &= b\left\{\exp\left[\frac{1}{2(1+\alpha)}\, a'(Z(Z'\Sigma^{-1}Z)^{-1}Z' - \alpha(\alpha+1)\Sigma)a\right]\right. \\ &\quad \left. + \frac{\alpha}{2(1+\alpha)}\, a'\Sigma a - 1\right\} \end{aligned} \tag{3.4}$$

The Bayes risk of $\hat{\delta}_{EB1}(X)$ is

$$R(\hat{\delta}_{EB1}(X),\theta) = b\left\{\exp\left[\frac{-\alpha}{2(\alpha+1)}\, a'\Sigma a\right] + \frac{\alpha}{2(\alpha+1)}\, a'\Sigma a - 1\right\} \tag{3.5}$$

The Bayes risk of $\hat{\delta}_{EB2}(X)$ is given by

$$\begin{aligned} R(\hat{\delta}_{EB2}(X),\theta) &= b\left\{\exp\left[\frac{1}{2(\alpha+1)}\, a'(D\Sigma - \alpha(\alpha+1)\Sigma)a\right]\right. \\ &\quad \left. + \frac{\alpha}{2(\alpha+1)}\, a'\Sigma a - 1\right\}. \end{aligned} \tag{3.6}$$

4 Comparisons

Comparing the risks of the Bayes estimators in the previous section, we have

(1) In the restricted model, under the LINEX loss function, the MLE of the location parameter θ is inadmissible. On the other hand, the Bayes estimator uniformly for every real vector-value a, dominates X. In this way, we have $\hat{\delta}_B(X) \succ X$; because easily it can be seen that for any scale parameter $b > 0$ and $\alpha > 0$ we have

$$R(\hat{\delta}_B(X), \theta) - R(X, \theta) = b\left\{\frac{\alpha}{2(1+\alpha)}\, a'\Sigma a - \exp\left[\frac{1}{2}\, a'\Sigma a\right]\right\} < 0 \tag{4.1}$$

(2) For any $b > 0$ and $\alpha > 0$, $\hat{\delta}_{EB}(X) \succ X$.

Because for a positive definite matrix $Z'\Sigma^{-1}Z$, we have $w = a'Z(Z'\Sigma^{-1}Z)^{-1}Z'a > 0$ (see Anderson 2003, p. 628);
taking

$$f_2(\alpha) = \exp\left[\frac{w}{2} \times \frac{1-\alpha-\alpha^2}{1+\alpha}\right] - \exp\left[\frac{w}{2}\right] + \frac{\alpha}{1+\alpha} \times \frac{w}{2},$$

we obtain

$$\frac{\partial f_2(\alpha)}{\partial \alpha} = \frac{-[(1+\alpha)^2+1]}{(1+\alpha)^2}\, \exp\left[\frac{w}{2} \times \frac{1-\alpha-\alpha^2}{1+\alpha}\right] + \frac{1}{(1+\alpha)^2} \times \frac{w}{2} < 0,$$

and also $\lim_{\alpha\to 0} f_2(\alpha) = 0$, $\lim_{\alpha\to\infty} f_2(\alpha) = \frac{w}{2} - e^{w/2} < 0$; thus $f_2(\alpha) < 0$, and by (3.1) and (3.4) we obtain

$$\begin{aligned} R(\hat{\delta}_{EB}(X), \theta) - R(X, \theta) &= b\left\{\exp\left[\frac{1}{2(\alpha+1)}\, a'(Z(Z'\Sigma Z)^{-1}Z' \right.\right. \\ &\quad \left.\left. - \alpha(\alpha+1)\Sigma)a\right] + \frac{\alpha}{2(\alpha+1)}\, a'\Sigma a - \exp\left[\frac{1}{2}a'\Sigma a\right]\right\} \\ &= \exp\left[\frac{w}{2}\left(\frac{1-\alpha-\alpha^2}{1+\alpha}\right)\right] - \exp\left[\frac{w}{2}\right] + \frac{\alpha}{\alpha+1}\,\frac{w}{2} < 0. \end{aligned}$$

(3) For every $a \in \Re^p$, $\hat{\delta}_{EB1}(X) \succ X$.

Using (3.1) and (3.5), the risk difference between $\hat{\delta}_{EB1}(X)$ and X is

$$\begin{aligned} f_1(\alpha) &= R(\hat{\delta}_{EB1}(X), \theta) - R(X, \theta) \\ &= b\left\{\exp\left[\frac{-\alpha}{2(\alpha+1)}\right] + \frac{\alpha}{2(\alpha+1)}\, a'\Sigma a - \exp\left[\frac{1}{2}\, a'\Sigma a\right]\right\} < 0, \end{aligned}$$

because for every $\alpha > 0$, $\frac{\alpha}{2(\alpha+1)} a'\Sigma a > 0$. One can find that $f_1(\alpha)$ is an increasing function with respect to α and that $\lim_{\alpha\to\infty} f_1(\alpha) < 0$; which gives the result.

(4) For $b > 0$ and $\alpha > 1$ we have $\hat{\delta}_{EB2}(X) \succ \hat{\delta}_{EB1}(X)$.

To see this, for any $a \in \Re^p$ let

$$\lambda_{\max} = \max_a \frac{a'D\Sigma a}{a'\Sigma a},$$

then by Theorem A.2.4. in (Anderson 2003), we have

$$\begin{aligned}\frac{a'D\Sigma a}{a'\Sigma a} &< \lambda_{\max}\\ &= ch_1(\Sigma^{-1}D\Sigma)\\ &= ch_1(D\Sigma\Sigma^{-1})\\ &= 1,\end{aligned}$$

where ch_1 is the biggest characteristic root.

Therefore, for $\alpha > 1$, we find $\alpha^2 > [a'D\Sigma a]/[a'\Sigma a]$. Consequently we obtain

$$\frac{a'D\Sigma a}{2(\alpha+1)} - \frac{\alpha}{2} a'\Sigma a < \frac{-\alpha}{2(\alpha+1)} a'\Sigma a,$$

and using (3.5) and (3.6) we have

$$R(\hat{\delta}_{EB2}(X), \theta) - R(\hat{\delta}_{EB1}(X), \theta) < 0$$

(5) If $X_1, X_2, \ldots, X_n$ are n iid r.v.s from $N_p(\theta, \Sigma)$ then all the above results are true for the sample mean. That is for $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ under the above conditions we have

$$\hat{\delta}_B(\underline{X}) \succ \bar{X}, \quad \hat{\delta}_{EB}(\underline{X}) \succ \bar{X},$$

where $\underline{X} = (X_1, X_2, \ldots, X_n)$.

Acknowledgments The authors would like to thank an anonymous referee and the Editor, Professor Ursula Gather for valuable comments that led to an improved version of the paper. Partial support from the "Ordered and Spatial Data Center of Excellence" of Ferdowsi University of Mashhad, Iran is acknowledged.

References

Anderson TW (2003) An introduction to multivariate statistical analysis, 3rd edn. Wiley, London

Efron B, Morris C (1972) Empirical Bayes on vector of observations: an extension of Stein's method. Biometrika 59:335–347

Efron B, Morris C (1976) Multivariate empirical Bayes and estimation of covariate matrices. Ann Math Stat 22–32

James W, Stein C (1961) Estimation with quadratic loss. In: Proceeding of fourth Berkdey symposium on mathematical statistics and probability, vol 1, pp 361–379

Saleh AKMdE (2006) Theory of preliminary test and stein-type estimation with applications. Wiley, New York

Srivastava MS, Saleh AKMdE (2005) Estimation of the mean vector of a multivariate normal distribution: subspace hypothesis. J Mult Anal 96:55–72