Journal of Mathematics and Statistics: 4(3): 145-147, 2008
ISSN: 1549-3644

Comparison of Estimators of Dispersion Matrix

Narges Abbasi
Department of Statistics, Payame Noor University, Shiraz Centre 71365-944, Shiraz, Iran

Abstract: Based on a sample, we considered the problem of estimating the dispersion matrix of a multivariate normal distribution with variance covariance matrix Σ. Empirical Bayes estimators and Haff estimators with their conditions, two proposed estimators of Σ, were the best affine equivariant estimators of dispersion matrix, which we compared them by three different loss functions.

Key words: Affine equivariant estimator, point estimation, loss function, risk function, Stein's estimator.

INTRODUCTION

Let $X_1, \dots X_N$ be i.i.d. $N_p(\mu, \Sigma)$, where μ and $\Sigma_{p\times p}$ are both unknown ($\mu \in R^p$ and Σ p.d.). We reduce the data set by sufficiency and concentrate only on $(\bar{X}, S)$, where $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \sim N_p(\mu, \frac{1}{N}\Sigma)$, $S = \sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})' \sim W_p(\Sigma, n)$ and n = N-1. The unbiased estimator of Σ is $\hat{\Sigma}_u = \frac{1}{n}S$. We evaluate the estimators by their functions or average loss functions. The loss functions for estimating Σ are:

$$L_1(\hat{\Sigma},\Sigma) = tr(\hat{\Sigma}\Sigma^{-1} - I_p)^2,$$
$$L_2(\hat{\Sigma},\Sigma) = tr(\hat{\Sigma}\Sigma^{-1}) - \ln(|\hat{\Sigma}\Sigma^{-1}|) - p,$$
$$L_3(\hat{\Sigma},\Sigma) = \exp(atr(\hat{\Sigma}\Sigma^{-1} - I_p)) - atr(\hat{\Sigma}\Sigma^{-1} - I_p) - 1,\ a \neq 0$$

By the group of affine transformation $(\bar{X}, S) \to (A\bar{X} + b, ASA')$ for nonsingular $A_{p\times p}$ be $b\in R^p$, the best affine equivariant estimators of Σ under L_1, L_2 and L_3 are respectively:

$$\hat{\Sigma}^1 = \frac{1}{N+p}S = c_1 S,$$
$$\hat{\Sigma}^2 = \frac{1}{n}S = c_2 S,$$
$$\hat{\Sigma}^3 = \frac{1}{2a}(1 - e^{-\frac{2ap}{np+2}})S = c_3 S$$

Empirical Bayes alternatives are derived which dominate all scalar multiples of the unbiased estimator. Empirical Bayes estimators have the form $\hat{\Sigma}_{EB} = b(S + ut(u)C)$ with $0 < b \le \frac{1}{n}, u = \frac{1}{trS^{-1}C}$, t(.) nonincreasing and C an arbitrary positive definite matrix. Note that the problem $(\Sigma, \hat{\Sigma}, L_i)$, i = 1, 2, 3 is invariant, so we can assume that C = I. To improved $\hat{\Sigma}^1$ and $\hat{\Sigma}^2$, Haff[2] obtained the following conditions on the estimators of the form:

$$\hat{\Sigma}_g^i = \hat{\Sigma}^i + g(S)C, \quad i = 1, 2$$

With $g(S) = c_i ut(u)$ i = 1,2, under L_1, L_2 respectively as:

- $0 < t < \frac{2(p-1)}{N-p+2}$, t is a constant
- t(u) is an absolutely continuous and nonincreasing function, $0 < t < \frac{2(p-1)}{n}$

Abbasi[1] derived some conditions for which $\hat{\Sigma}^3$ is dominated by $\hat{\Sigma}_g^3$ under loss L_3. Pal and Elfessi[3] proposed that in the expression of S, used the James-Stein structure instead of $\bar{X}$. They start with:

$$\hat{\Sigma}_{c,\alpha}^i = \hat{\Sigma}^i + \frac{c}{(\bar{X}'S^{-1}\bar{X})}\bar{X}\,\bar{X}'$$

As a new estimator of Σ. Motivated by[4], when p≥2, one can also use $\bar{X}$ to get improvements but such typical improved estimators have one undesirable property, they are nonanalytic and hence inadmissible. This estimator is scale equivariant and uses both $\bar{X}$ and S. For p≥2, if $0 < c \le \frac{2(p-1)}{(N+p)(N-p+2)}$, then $\hat{\Sigma}_{c,1}^1$ is uniformly better than $\hat{\Sigma}^1$ under loss L_1 and if $0 < c \le \frac{p-1}{(N-1)(N-p)}$, $\hat{\Sigma}_{c,1}^2$ is uniformly better than $\hat{\Sigma}^2$ under loss L_2.

Tsukuma and Konno[5] considered the problem of estimating the precision matrix of a multivariate normal distribution model with respect to a quadratic loss function. Furthermore, a numerical study was undertaken to compare the properties of a collection of alternatives to the "unbiased" estimator of the discriminate coefficients.

In this research, under some assumptions, the comparison of $\hat{\Sigma}^i_g$ and $\hat{\Sigma}^1_{c,1}$ for i = 1, 2, 3 is considered.

MATERIALS AND METHODS

It is obvious that the role of c_i in changing of the amount of risk function is essential. With determining the boundaries for c_i, two groups of estimators are compared under three loss functions.

RESULTS AND DISCUSSION

Comparison under L_1: It is well known that, given $\bar{X}$, $\frac{\bar{X}'\Sigma^{-1}\bar{X}}{\bar{X}'S^{-1}\bar{X}} \sim \chi^2_{N-p}$ which is free from $\bar{X}$ So, the risk function for $\hat{\Sigma}^1_{c,1}$ under L_1 is:

$$R_1(\hat{\Sigma}^1_{c,1},\Sigma) = R_1(\hat{\Sigma}^1,\Sigma) + c^2(N-p)(N-p+2) - 2c\tfrac{(N-p)(p-1)}{N+p}$$

and the risk function for $\hat{\Sigma}^1_g$ under L_1 is:

$$R_1(\hat{\Sigma}^1_g,\Sigma) = R_1(\hat{\Sigma}^1,\Sigma) + \alpha_1(\Sigma)$$

Where:

$$\alpha_1(\Sigma) = E(\tfrac{2}{N+p}g(S)\mathrm{tr}(S\Sigma^{-2}) - 2g(S)\mathrm{tr}(\Sigma^{-1}) + g^2(S)\mathrm{tr}(\Sigma^{-2})).$$

Haff[2] obtained an unbiased estimator for $\alpha_1(\Sigma)$ and showed that:

$$\alpha_1(\Sigma) \le (\tfrac{1}{N+p})^2(-2(N-p)(p-1)t + (N-p+2)(N-p)t^2)$$

for condition (I), $\alpha_1(\Sigma) \le 0$.

Now, we consider the risk difference:

$$\begin{aligned} RD_1 &= R_1(\hat{\Sigma}^1_{c,1},\Sigma) - R_1(\hat{\Sigma}^1_g,\Sigma) \\ &= c^2(N-p)(N-p+2) - \frac{2c(N-p)(p-1)}{N+p} - \alpha_1(\Sigma) \\ &\ge (N-p)(c^2(N-p+2) - \frac{2c(p-1)}{N+p} + \\ &(\frac{1}{N+p})^2(2(p-1)t - (N-p+2)t^2)). \end{aligned}$$

So $RD_1 \ge 0$ and it implies the following theorem.

Theorem 1: Under loss function L_1, the estimator $\hat{\Sigma}^1_g$ dominates $\hat{\Sigma}^1_{c,1}$ if:

$$0 < t < \frac{p-1}{N-p+2}, \quad (0 < c < \frac{t}{N+p} \text{ or } c > \frac{2(p-1)(N-p+2)t}{(N+p)(N-p+2)})$$

or

$$\frac{p-1}{N-p+2} < t < \frac{2(p-1)}{N-p+2}, \quad (0 < c < \frac{2(p-1)-(N-p+2)t}{(N+p)(N-p+2)} \text{ or } c > \frac{t}{N+p}).$$

Comparison under L_2: The risk function for $\hat{\Sigma}^2_{c,1}$ under L_2 is:

$$\begin{aligned} R_2(\hat{\Sigma}^2_{c,1},\Sigma) &= R_2(\hat{\Sigma}^2,\Sigma) + c(N-p) - \ln(1+cn) \\ &\ge R_2(\hat{\Sigma}^2,\Sigma) - (p-1)c, \end{aligned}$$

(since $\ln(1+x) < x$ for $x > 0$) and:

$$\begin{aligned} R_2(\hat{\Sigma}^2_g,\Sigma) &= R_2(\hat{\Sigma}^2,\Sigma) + \alpha_2(\Sigma) \\ &\le R_2(\hat{\Sigma}^2,\Sigma) + E(\tfrac{1}{n}t(u)(\tfrac{n}{2}t(u) - (p-1))) \end{aligned}$$

where, $\alpha_2(\Sigma) = E(g(S)\mathrm{tr}\Sigma^{-1} - \ln|I + \frac{1}{n}g(S)S^{-1}|)$. Therefore:

$$RD_2 \le E(\tfrac{1}{n}t(u)(\tfrac{n}{2}t(u) - (p-1))) - (p-1)c$$

For $RD_2 \ge 0$, we have the following theorem.

Theorem 2: Under loss function L_2, with condition (II), the estimator $\hat{\Sigma}^2_g$ dominates $\hat{\Sigma}^2_{c,1}$ if $0 < c < \frac{1}{n}t(u)(1 - \frac{n}{2(p-1)}t(u))$.

Comparison under L_3: The risk function for $\hat{\Sigma}^3_{c,1}$ under loss L_3 with a > 0 is:

$$\begin{aligned} R_3(\hat{\Sigma}^3_{c,1},\Sigma) &= E\{\exp(a\,\mathrm{tr}(\hat{\Sigma}^3_{c,1}\Sigma^{-1} - I)) - a\,\mathrm{tr}(\hat{\Sigma}^3_{c,1}\Sigma^{-1} - I) - 1\} \\ &= e^{-ap}(1-2ac_3)^{-\frac{np}{2}}(1-2ac)^{-\frac{N-p}{2}} - ac_3np \\ &-ac(N-p) + ap - 1, \end{aligned}$$

and the risk function for $\hat{\Sigma}^3_g$ is:

$$R_3(\hat{\Sigma}_g^3,\Sigma) = E\{\exp(a\,tr(\hat{\Sigma}_g^3\Sigma^{-1}-I)) - a\,tr(\hat{\Sigma}_g^3\Sigma^{-1}-I)-1\}$$
$$\geq E\{\exp(a\,tr(\hat{\Sigma}^3\Sigma^{-1}-I)) - a\,tr(\hat{\Sigma}_g^3\Sigma^{-1}-I)-1\}$$
$$\geq e^{-ap}(1-2ac_3)^{-\frac{np}{2}} - ac_3np - aE(g(S)tr\Sigma^{-1}) + ap - 1$$
$$= e^{-ap}(1-2ac_3)^{-\frac{np}{2}} - ac_3np$$
$$-ac_3E((n-p-1)t(u) + 2(ut'(u) + t(u)\frac{trS^{-2}}{(trS^{-1})^2}))$$
$$+ap-1$$

Therefore:

$$RD_3 \leq e^{\frac{-2ap}{np+2}}((1-2ac)^{-\frac{N-p}{2}}-1) - ac(N-p)$$
$$+ac_3E((n-p-1)t(u) + 2(ut'(u) + t(u)\frac{trS^{-2}}{(trS^{-1})^2}))$$

Since $\frac{trS^{-2}}{(trS^{-1})^2} \leq 1$,

$$RD_3 \leq e^{\frac{-2ap}{np+2}}((1-2ac)^{-\frac{N-p}{2}}-1) - ac(N-p)$$
$$+ac_3E((n-p-1)t(u) + 2(ut'(u) + t(u)))$$

$RD_3 \leq 0$, if we have the following theorem.

Theorem 3: Under loss function L_3, if t(u) is an absolutely continuous and nonincreasing function and:

$$0 < t(u) \leq \frac{ac(N-p)+(1-2ac_3)(1-(1-2ac)^{-\frac{N-p}{2}})}{a(N-p)c_3}$$

then for a>0 the estimator $\hat{\Sigma}_{c,1}^3$ dominates $\hat{\Sigma}_g^3$ and for a<0 the estimator $\hat{\Sigma}_g^3$ dominates $\hat{\Sigma}_{c,1}^3$.

CONCLUSION

The finding of this article suggests that with the changing of t(u) and determining it for special cases, there will be new characteristics for the estimator, $\hat{\Sigma}_g^i$.

REFERENCES

1. Abbasi, N.,1999.Estimating the Normal Dispersion Matrix and the Precision Matrix, Ph. D. thesis, Shiraz University. dbase.irandoc.ac.ir/00292/00292339.htm.
2. Haff, L.R., 1980. Empirical bayes estimation of the multivariate normal covariance matrix. Ann. Stat., 8:586-597. http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.aos/1176345010.
3. Pal, N. and A. Elfessi, 1995. Improved estimation of a multivariate normal mean vector and the dispersion matrix, how one affects the other. Indian J.Stat.,57:267-286. http://www.jstor.org/pss/25051052.
4. Stein, C., 1964. Inadmissibility of the usual estimator for the normal distribution with unknown mean. Ann. Inst. Stat. Math., 16: 155-160. DOI: 10.1007/BF02868569.
5. Tuskuma, H. and Y. Konno, 2006. On improved estimation of normal precision matrix and discriminate coefficients. J. Multivariate Anal., 97: 1477-1500. DOI: 10.1016/j.jmva.2005.11.006.