

ASYMPTOTIC RISK COMPARISON OF IMPROVED ESTIMATORS FOR NORMAL COVARIANCE MATRIX

By

Nariaki SUGIURA and Masahiro FUJIMOTO

Asymptotic risks of the empirical Bayes estimators $\hat{\Sigma}_H$ by Haff [5] for a covariance matrix Σ in a p -dimensional normal distribution are computed and compared with that of James and Stein's minimax estimators $\hat{\Sigma}_{JS}$. For $p \geq 6$, it is shown that $\hat{\Sigma}_{JS}$ are always better than $\hat{\Sigma}_H$ asymptotically, though the leading terms are the same. New estimators which dominate $\hat{\Sigma}_{JS}$ for some Σ in any p asymptotically are proposed. Some numerical comparisons are given. Exact risks for ordinary estimators $\hat{\Sigma}_O$ and minimax estimators $\hat{\Sigma}_{JS}$ are also computed and compared with asymptotic ones for which the approximations are shown to be excellent.

1. Introduction

Let S have a Wishart distribution with unknown scale matrix Σ and n degrees of freedom, for which we shall write $S: W_p(n, \Sigma)$ and assume $n > p + 1$. Let $\hat{\Sigma}$ be an estimator of Σ . The loss function is taken to be

$$(1.1) \quad L_1(\hat{\Sigma}, \Sigma) = \text{tr } \hat{\Sigma} \Sigma^{-1} - \log |\hat{\Sigma} \Sigma^{-1}| - p$$

or

$$(1.2) \quad L_2(\hat{\Sigma}, \Sigma) = \frac{1}{2} \text{tr}(\hat{\Sigma} \Sigma^{-1} - I)^2.$$

The L_1 loss is equivalent to the likelihood ratio statistic for testing the hypothesis $\Sigma = \Sigma_0$ against all alternatives. The L_2 loss can also be used as a test statistic for the same problem as in Nagao [10]. The factor $1/2$ in the L_2 loss is not essential. However we wish to retain it, since L_1 loss tends to $\text{tr}(\hat{\Sigma} \Sigma^{-1} - I)^2/2$, when $\hat{\Sigma}$ is close to Σ . The risk function is given by $R_i(\hat{\Sigma}, \Sigma) = E[L_i(\hat{\Sigma}, \Sigma)]$ for $i=1$ or 2 . Haff [5] proved that among the scalar multiples of S , the best estimator under L_1 is $\hat{\Sigma}_O^{(S)} = S/n$ and that under L_2 it is given by $\hat{\Sigma}_O^{(S)} = S/(n+p+1)$, which we call ordinary estimators. Then he considered the posterior mean of Σ for a prior distribution $W_p[n', (\gamma C)^{-1}]$ for Σ^{-1} with unknown scalar $\gamma > 0$ and known p. d. matrix