Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press, New York and London
Vol. 40, No. 1, January 1992
Printed in Belgium

On a Shrinkage Estimator of a Normal Common Mean Vector

K. KRISHNAMOORTHY*

Department of Statistics, Temple University, Philadelphia

Communicated by the Editors

The problem of estimating the $p \times 1$ mean vector θ based on two independent normal vectors $Y_1 \sim N_p(\theta, \sigma^2 I)$ and $Y_2 \sim N_p(\theta, \xi\sigma^2 I)$ is considered. For $p \geqslant 3$, when ξ and σ^2 are unknown, it was shown by George (1991, *Ann. Statist.*) that under certain conditions estimators of the form $\delta_\eta = \eta Y_1 + (1-\eta) Y_2$, where η is a fixed number in (0, 1), are uniformly dominated by a shrinkage estimator under the squared error loss. In this paper, George's result is improved by obtaining a simpler and better condition for the domination. © 1992 Academic Press, Inc.

1. INTRODUCTION

Let Y_1 be an observation from a p-variate normal population with mean θ and covariance matrix $\sigma^2 I$, $N_p(\theta, \sigma^2 I)$. Let Y_2 be an observation from $N_p(\theta, \xi\sigma^2 I)$. Assume that σ^2, ξ are unknown positive scalars and Y_1 and Y_2 are independent. Let $\Psi \equiv \{\psi : \theta \in R^p,\ \xi > 0,\ \sigma^2 > 0\}$ denote the parameter space. In this situation George [2] has considered the problem of estimating θ based on (Y_1, Y_2) under the risk criterion of expected squared error loss,

$$R(\psi, \delta) = E_\psi \|\delta - \theta\|^2. \tag{1.1}$$

Under this criterion, it has been shown in George [2] that the estimator

$$\delta_\eta = \eta Y_1 + (1-\eta) Y_2 \tag{1.2}$$

is dominated by the shrinkage estimator

$$\delta_\eta^c = \left(1 - c\,\frac{\|Y_1 - Y_2\|^2}{\|\delta_\eta\|^2}\right)\delta_\eta, \tag{1.3}$$

Received December 4, 1990; revised April 16, 1991.
AMS 1980 subject classifications: primary 62F15; secondary 62F10, 62F25.
Key words and phrases: loss function, shrinkage estimator, Poisson distribution.

* Current address: Department of Mathematics and Statistics, University of South Alabama, FCS 3, Mobile, AL 36688.

0047-259X/92 $3.00

where η is a fixed number in $(0, 1)$, $\kappa = \min(\eta^2, (1-\eta)^2)$, $0 < c < 2\kappa/13$ for $p = 3$ and $0 < c < 2(p-2)\kappa/(p+8)$ for $p \geqslant 4$. Based on strong simulation evidence, George [2] has conjectured that the upper bound for c can be increased to $2\kappa/6$ for $p = 3$ and to $2(p-2)\kappa/(p+2)$ for $p \geqslant 4$.

In the following section we show that, for δ_η^c to dominate δ_η, the upper bound for c can in fact be increased to $2(p-2)\kappa/(p+2)$ not only for $p \geqslant 4$ but also for $p = 3$.

We note that although some expectations that are needed to prove Theorem 2.1 are evaluated in George [2], we derive them in Section 2 in a relatively easy manner using some standard methods. Also the present form of these expectations simplified the proof of Theorem 2.1 to some extent.

2. Main Results

We now state the main results of this paper in the following theorem.

THEOREM 2.1. *Let* $\kappa \equiv \min(\eta^2, (1-\eta)^2)$. *Then for* $p \geqslant 3$, $\eta \in (0, 1)$ *and* $0 < c < 2(p-2)\kappa/(p+2)$, *the estimator* δ_η^c *in* (1.3) *uniformly dominates* δ_η *under the risk criterion* (1.1).

We need the following lemmas to prove the theorem.

LEMMA 2.1. *For a Poisson random variable* Z *with mean* λ, *let* $\phi(q) \equiv E(q+2Z)^{-1}$ *for* $q \neq 0, -2, -4, \ldots$, *then*

(i) $(q+2\lambda)\,\phi(q) \geqslant 1$,

(ii) $2\lambda\phi(q+2) = 1 - q\phi(q)$, *for* $q \neq 0, -2, -4, \ldots$,

(iii) $2\lambda\phi(q+2) \leqslant 1 - q\phi(q+2)$, *for* $q \geqslant 0$,

(iv) $5\alpha[\phi(-1) - 3\phi(1)] + 3\alpha^2[\phi(1) - 2\phi(-1) + \phi(-3)] < 0$, *for* $0 \leqslant \alpha \leqslant 1$.

Proof. (i) follows from Jensen's inequality. For proofs of (ii) and (iii), see George [2].

(iv) As $(2z-1)^{-1} - 3(2z+1)^{-1} \leqslant 0$ for $z = 0, 1, 2, \ldots$, and $\alpha \geqslant \alpha^2$, we have

$$\begin{aligned} &5\alpha[(2z-1)^{-1} - 3(2z+1)^{-1}] + 3\alpha^2[(2z+1)^{-1} - 2(2z-1)^{-1} + (2z-3)^{-1}] \\ &\quad \leqslant \alpha^2[5(4-4z)(2z-3) + 24]/[(2z+1)(2z-1)(2z-3)] \\ &\quad = -4\alpha^2\gamma(z), \end{aligned}$$

where $\gamma(z) = (10z^2 - 25z + 9)/[(2z-1)(2z+1)(2z-3)]$. Thus, to prove (iv), it suffices to show that $E(\gamma(Z))>0$. Observing that $\gamma(z)>0$ for nonnegative integers $z\neq 2$, we prove that

$$E(\gamma(Z))=\sum_{z\neq 2,3}\gamma(z)e^{-\lambda}\lambda^{z}/z!+[(8\lambda/7)-1]e^{-\lambda}\lambda^{2}/30>0$$

for $\lambda>7/8$. Similarly, adding the terms corresponding to $z=0$ and $z=2$, we can show that $E(\gamma(Z))>0$ for $0<\lambda\leqslant 7/8$.

LEMMA 2.2. *Let U and V be $p\times 1$ random vectors such that $(U', V')'\sim N_{2p}((0', \mu')', \Sigma_\beta)$ with $\Sigma_\beta=\begin{pmatrix} I & \beta I\\ \beta I & I\end{pmatrix}$ and $-1<\beta<1$. For a Poisson random variable Z with mean $\lambda=\|\mu\|^2/2$, as in Lemma 2.1, let $\phi(q)\equiv E(q+2Z)^{-1}$. Then, for $p\geqslant 3$,*

(i) $E(\|U\|^2\|V\|^{-2}=[\beta^2(p-4+2\lambda)+p]\,\phi(p-2)-\beta^2$

(ii) $E(\|U\|^4\|V\|^{-2})=[(p+4)+\beta^2(p-6+2\lambda)]\,E(\|U\|^2\|V\|^{-2})-[2(2\beta^4\lambda-\beta^2p+p)\,\phi(p-2)+\beta^2(p-2\beta^2)]$.

(iii) $E(U'V\|V\|^{-2})=\beta(p-2)\,\phi(p-2)$.

Proof. All these expectations can be evaluated using the well-known identity that

$$E\|V\|^{-2}=E(p-2+2Z)^{-1}. \tag{2.1}$$

Let $f(U, V;\mu, \Sigma)=(2\pi)^{-p}\exp\{-1/2[U', (V-\mu)']\,\Sigma^{-1}[U', (V-\mu)']'\}$.

(i) First we note that for a real t,

$$\begin{aligned} E(\|U\|^2\,\|V\|^{-2}) &= -\frac{\partial}{\partial t}\left\{(1-\beta^2)^{-p/2}\int\|V\|^{-2}\,e^{-tU'U}f(U, V;\mu, \Sigma_\beta)\,d_U\,d_V\right\}\Big|_{t=0}\\ &= -\frac{\partial}{\partial t}\left\{(1-\beta^2)^{-p/2}\int\|V\|^{-2}f(U, V;\mu, \Sigma_{\beta,t})\,d_U\,d_V\right\}\Big|_{t=0}, \end{aligned} \tag{2.2}$$

where

$$\Sigma_{\beta,t}=\begin{pmatrix}(1+2t)^{-1}I & \beta(1+2t)^{-1}I\\ \beta(1+2t)^{-1}I & (1+2t(1-\beta^2))(1+2t)^{-1}I\end{pmatrix}.$$

Noting that $|\Sigma_{\beta,t}|=(1-\beta^2)^p\,(1+2t)^{-p}$, it follows from (2.2) that

$$E(\|U\|^2\,\|V\|^{-2})=-\frac{\partial}{\partial t}\{(1+2t)^{-p/2+1}\,(1+2t(1-\beta^2))^{-1}\,E\|V_t\|^2\}|_{t=0}, \tag{2.3}$$

where $V_t \sim N_p(\mu_t, I)$ and $\mu_t = (1+2t)^{1/2}(1+2t(1-\beta^2))^{-1/2}\mu$. Thus, we obtain

$$E(\|U\|^2\|V\|^{-2})$$

$$= -\frac{\partial}{\partial t}\{(1+2t)^{-p/2+1}(1+2t(1-\beta^2))^{-1}E(p-2+2Z_t)^{-1}\}|_{t=0}$$

$$= (p-2\beta^2+2\beta^2\lambda)E(p-2+2Z)^{-1} - 2\beta^2 E(Z(p-2+2Z)^{-1}),$$

where Z_t in the first equality is a Poisson random variable with mean $\lambda_t = \mu_t'\mu_t/2$. Now, using the fact that $E(2Z(p-2+2Z)^{-1}) = 1-(p-2)\phi(p-2)$, we prove (i).

(ii) As in (i), we can write

$$E(\|U\|^4\|V\|^{-2})$$

$$= -\frac{\partial}{\partial t}\left\{(1-\beta^2)^{-p/2}\int\|U\|^2\|V\|^{-2}f(U,V;\mu,\Sigma_{\beta,t})\,d_U\,d_V\right\}\Big|_{t=0}$$

$$= -\frac{\partial}{\partial t}\{(1+2t)^{-p/2}(1+2t(1-\beta^2))^{-1}E(\|U_t\|^2\|V_t\|^{-2})\}|_{t=0}, \tag{2.4}$$

where $(U_t', V_t')' \sim N_{2p}((0', \mu_t')', \Sigma_{\beta_t})$, and Σ_{β_t} equal to Σ_β with β replaced by $\beta_t = \beta(1+2t(1-\beta^2))^{-1/2}$. Thus, using (i) in (2.4), taking derivative at $t=0$, and after some simplification, we get (ii).

(iii) $E(U'V\|V\|^{-2}) = EE((U'V\|V\|^{-2})\,|\,V) = \beta E((V-\mu)'V\|V\|^{-2}) = \beta(1-E(\mu'V\|V\|^{-2})$ and $E(\mu'V\|V\|^{-2}) = \mu'(\partial/\partial\mu)E\|V\|^{-2} + \|\mu\|^2 E\|V\|^{-2} = E(2Z(p-2+2Z)^{-1})$. Thus, we prove (iii).

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. Using the Stein's [3] identity that $E(X-\mu)h(X) = \sigma^2 Eh'(X)$ for $X \sim N(\mu, \sigma^2)$ and changing the variables, George [2] has expressed the risk difference as

$$R(\psi, \delta_n^c) - R(\psi, \delta_n) = c\sigma_n^2(cA_1 - 2(p-2)A_2 - 4\sigma_n^{-2}A_3), \tag{2.5}$$

where

$$A_1=(1+\xi)^2\,\sigma^4\sigma_\eta^{-4}E(\|U\|^4\,\|V\|^{-2}),$$
$$A_2=(1+\xi)\,\sigma^2\sigma_\eta^{-2}E(\|U\|^2\,\|V\|^{-2}),$$
$$A_3=(1+\xi)\,\sigma^2\beta E(U'V\|V\|^{-2}),$$
$$\begin{pmatrix}U\\V\end{pmatrix}\sim N_{2p}\left(\begin{pmatrix}0\\\mu\end{pmatrix},\begin{pmatrix}I&\beta I\\\beta I&I\end{pmatrix}\right),$$
$$\beta=\frac{(\eta-(1-\eta)\xi)\sigma}{\sigma_\eta\sqrt{(1+\xi)}},\qquad \mu=\theta/\sigma_\eta,\qquad \sigma_\eta^2=(\eta^2+(1-\eta)^2\xi)\sigma^2.$$

In terms of these expectations, (2.5) may be expressed as

$$\begin{aligned}&R(\psi,\delta_\eta^c)-R(\psi,\delta_\eta)\\&\quad=c(1+\xi)\,\sigma^2[c(1+\xi)\,\sigma^2\sigma_\eta^{-2}E(\|U\|^4\,\|V\|^{-2})\\&\qquad-2(p-2)\,E(\|U\|^2\,\|V\|^{-2})-4\beta E(U'V\|V\|^{-2})].\end{aligned}\tag{2.6}$$

As $(1+\xi)\sigma^2\sigma_\eta^{-2}\leqslant\kappa^{-1}$, for $0<c<2(p-2)\kappa/(p+2)$, it can be easily seen that $R(\psi,\delta_\eta^c)-R(\psi,\delta_\eta)<0$ if

$$\begin{aligned}g(\eta,\xi,\theta)=&E(\|U\|^4\,\|V\|^{-2})-(p+2)\,E(\|U\|^2\,\|V\|^{-2})\\&-2\beta(p-2)^{-1}\,(p+2)\,E(U'V\|V\|^{-2})\leqslant 0\end{aligned}$$

for all η, ξ, and θ. Using Lemma 2.2 and setting $\alpha=\beta^2$, $g(\eta,\xi,\theta)$ can be written as

$$\begin{aligned}g(\eta,\xi,\theta)=&[2+\alpha(p-6+2\lambda)]\,E\,\|U\|^2\,\|V\|^{-2}-\alpha(p-2\alpha)\\&-2(2\alpha^2\lambda-\alpha p+p)\,\phi(p-2)-2\alpha(p+2)\,\phi(p-2).\end{aligned}\tag{2.7}$$

We now show that $g(\eta,\xi,\theta)\leqslant 0$ for all η, ξ, and θ, separately for $p\geqslant 4$ and $p=3$.

Let $p\geqslant 4$. We first note that $2+\alpha(p-6+2\lambda)\geqslant 0$ as $0\leqslant\alpha<1$. Using Lemma 2.1(iii) in Lemma 2.2(i), we obtain $E\|U\|^2\,\|V\|^{-2}\leqslant p\phi(p-2)$. Using this inequality in (2.7) and, after some minor simplification, we obtain

$$\begin{aligned}g(\eta,\xi,\theta)&\leqslant\alpha(p-4+2\lambda)\,p\phi(p-2)-4\alpha^2\lambda\phi(p-2)-\alpha p\\&\quad+2\alpha^2-2\alpha(p+2)\,\phi(p-2)\\&\leqslant -4\alpha^2\lambda\phi(p-2)+2\alpha^2-2\alpha^2(p+2)\,\phi(p-2)\\&=2\alpha^2-8\alpha^2\phi(p-2)-2\alpha^2(p-2+2\lambda)\,\phi(p-2).\end{aligned}\tag{2.8}$$

The second inequality in the above expression follows from Lemma 2.1(iii) and the relation $\alpha \geqslant \alpha^2$. Now the desired result follows from (2.8) and the fact that $(p-2+2\lambda)\,\phi(p-2) \geqslant 1$ (Lemma 2.1(i)).

For $p=3$, using Lemma 2.1(ii) in Lemma 2.2(i), we obtain $E\|U\|^2\,\|V\|^{-2} = \alpha\phi(-1) - \alpha\phi(1) + 3\phi(1)$. Substituting this equality in (2.7) and, after some simplification, we can write

$$g(\eta, \xi, \theta) = 2\alpha\phi(-1) - 15\alpha\phi(1) + 3\alpha^2\phi(1) - 3\alpha^2\phi(-1) - 3\alpha + 2\alpha^2 - 6\alpha^2\lambda\phi(1) + 6\alpha\lambda\phi(1) + 2\alpha^2\lambda\phi(-1). \tag{2.9}$$

Again from Lemma 2.1(ii), we obtain $2\lambda\phi(1) = 1 + \phi(-1)$ and $2\lambda\phi(-1) = 1 + 3\phi(-3)$. Uing these identities in (2.9), we can express

$$g(\eta, \xi, \theta) = 5\alpha[\phi(-1) - 3\phi(1)] + 3\alpha^2[\phi(1) - 2\phi(-1) + \phi(-3)].$$

Now, Lemma 2.1(iv) completes the proof.

Remark 2.1. Noting that α becomes zero when $\eta = \xi/(1+\xi)$, it can be easily verified that $R(\psi, \delta_\eta^c) - R(\psi, \delta_\eta) = 0$ when $\xi = 1$ and $c = 2(p-2)\kappa/(p+2)$. This implies that, for δ_η^c to dominate δ_η, the least upper bound for c in δ_η^c is $2(p-2)\kappa/(p+2)$.

Remark 2.2. For $p \geqslant 3$, it follows from (2.5) that an optimal choice for c in δ_η^c is $(p-2)\kappa/(p+2)$.

Remark 2.3. The positive part version of δ_η^c, that is, $(1 - c\|Y_1 - Y_2\|^2\,\|\delta_\eta\|^{-2})^+\,\delta_\eta$, where $a^+ = \max(0, a)$ improves δ_η^c uniformly. The proof, noting the fact that $g(Y_1, Y_2) = (1 - c\|Y_1 - Y_2\|^2\,\|\delta_\eta\|^{-2}) = g(\Gamma Y_1, \Gamma Y_2)$ for any orthogonal matrix Γ, is similar to the one given for one sample James–Stein estimation problem (for example, see Anderson [1, Lemma 3.5.2, p. 91]).

Acknowledgment

The author is grateful to the referee for his valuable comments and suggestions.

References

[1] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] George, E. I. (1991). Shrinkage domination in a multivariate common mean problem. *Ann. Statist.* **19** 952–960.

[3] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium