Proceedings of the Second Asian Mathematical Conference, 1995

SOME PROBLEMS ON MINIMAX ESTIMATION

SUBANAR
Department of Mathematics
Faculty of Science and Mathematics
Gadjah Mada University, Yogyakarta, Indonesia

We are concerned with minimax estimation related to various restrictions on the parameter sets. As explicit minimax estimators can only be found for small classes of problems, we solve our problems asymptotically. The key idea here is the relation between the minimax risk and the Bayes risk. In this paper, we show how to create an appropriate Bayes estimator in which the Bayes risk coincides asymptotically with the minimax risk.

1. Introduction

Suppose we are given $X_1, X_2, \dots, X_p$ independent normal random variables where each X_i has mean θ_i and variance 1. Let $\tilde{\theta} = (\theta_1, \dots, \theta_p)'$ and $\tilde{x} = (x_1, \dots, x_p)'$. We assume that $\tilde{\theta}$ varies in a parameter space Ω where Ω is a sphere or a cube.

We wish to estimate $\tilde{\theta}$ using an estimator $\delta(\tilde{x}) = \big(\delta_1(\tilde{x}), \delta_2(\tilde{x}), \dots, \delta_p(\tilde{x})\big)'$ under a general quadratic loss function $L(\tilde{\delta}, \tilde{\theta}) = (\tilde{\delta} - \tilde{\theta})' Q (\tilde{\delta} - \tilde{\theta})$, where Q is a known positive definite matrix, and we will use the minimax principle to evaluate estimates. An estimator δ^* is minimax if

$$\sup_{\theta \in \Omega} R(\delta^*, \theta) = \inf_{\delta} \sup_{\theta \in \Omega} R(\delta, \theta),$$

where $R(\delta, \theta) = EL(\delta, \theta)$ is the risk of $\delta(\tilde{x})$.

It is generally know that in this case the explicit minimax estimator is not easy to find, and the problem will need to be treated asymptotically. Our concern here is to investigate the asymptotic behavior of the minimax risk up to a second order term and to obtain a second order asymptotic minimax estimator.

For Poisson data, considerable literature has recently become available, relating such minimax risks to the size and shape of the constraints and to the structure of the loss function, e.g. Clevenson & Zidek [6], or Johnstone & Mac Gibbon [8].

Under the squared error loss and when the parameter space is restricted to a sphere of radius m, Bickel [4] gave the asymptotic form of the minimax risk together with an estimator which is asymptotically second order minimax as m tends to infinity. In section 2 we generalize Bickel's result to the case of a weighted quadratic loss function (Q is a diagonal matrix) and the case where the parameter space is a p-dimensional cube or an infinite dimensional cube. Our main result in this extension is a relationship between the Fisher information and the minimum eigenvalue of the Dirichlet problem for an elliptic partial differential equation in the parameter space. Finally, in section 3 using a suitable transformation from the decomposition $C'QC = D = \operatorname{diag}(\lambda_1, \dots, \lambda_p)$, the problem of estimating $\tilde{\theta}$ where $\tilde{\theta}$ varies in a sphere or a p-dimensional cube under a general quadratic loss function $L(\tilde{\delta}, \tilde{\theta}) = (\tilde{\delta} - \tilde{\theta})'Q(\tilde{\delta} - \tilde{\theta})$ can be simplified to the problem of a weighted quadratic loss function.

2. Minimax estimation for a weighted quadratic loss function

Let $\tilde{X}$ be a p-variate normal distribution with mean vector $\tilde{\theta} = (\theta_1, \dots, \theta_p)'$ with a covariance matrix I. First we assume that

$$L(\tilde{\delta}, \tilde{\theta}) = \sum_{i=1}^{p} \lambda_i (\delta_i(\tilde{x}) - \theta_i)^2$$

where $\lambda_i > 0$, $i = 1, 2, \dots, p$. We denote the risk of an estimator $\delta(\tilde{x})$ by $R(\tilde{\delta}, \tilde{\theta}) = E_\theta L(\tilde{\delta}, \tilde{\theta})$. The Bayes risk with respect to a probability distribution $G(\theta)$ is given by $r(G, \delta) = \int R(\tilde{\delta}, \tilde{\theta})\, dG(\theta)$. Denote the Bayes estimator, i.e. the estimator which minimizes $r(G, \delta)$ by $\delta(*, G)$, and its corresponding Bayes risk $r(G, \delta(*, G))$ by $r(G)$. The minimax risk for estimating $\tilde{\theta}$, provided that $\tilde{\theta} \in \Omega$, is given by $\rho_p(m) = \inf_{\delta \in \mathcal{D}} \sup_{\theta \in \Omega} R(\theta, \delta)$.

Preliminary lemmas. Let $\psi(\tilde{x}) = \big(\psi_1(\tilde{x}), \psi_2(\tilde{x}), \dots, \psi_p(\tilde{x})\big)'$ and $\delta(\tilde{x}) = \tilde{x} - \psi(\tilde{x})$. By Theorem 1 of Stein [9] we have

$$R(\delta, \theta) = \sum_{i=1}^{p} \lambda_i + \sum_{i=1}^{p} \lambda_i E\big(\psi_i^2(\tilde{x})\big) - 2 \sum_{i=1}^{p} \lambda_i E\left[\frac{\partial}{\partial x_i} \psi_i(\tilde{x})\right]. \tag{2.1}$$

Hence the Bayes risk with respect to a proper prior $G(\tilde{\theta})$ is equal to

$$r(G, \delta) = \sum_{i=1}^{p} \lambda_i + \sum_{i=1}^{p} \lambda_i \int_{\mathbb{R}^p} \psi_i^2(\tilde{x}) f_G(\tilde{x}) dx - 2 \sum_{i=1}^{p} \lambda_i \int_{\mathbb{R}^p} \frac{\partial \psi_i(\tilde{x})}{\partial x_i} f_G(\tilde{x}) d\tilde{x} \tag{2.2}$$

where

$$f_G(\tilde{x}) = \int_{\Omega_m} \frac{e^{-\frac{1}{2}\|\tilde{x} - \tilde{\theta}\|^2}}{(\sqrt{2\pi})^{p/2}}\, dG(\tilde{\theta}).$$

From formula (3.3) of [9] with
estimator is the posterior mea

$\delta(\tilde{x}$

Moreover, $\delta(\tilde{x}, G)$ is also the B
$\sum_{i=1}^{p} \lambda_i (G_i(\tilde{x}) - \theta_i)^2$ ([3], page
risk becomes

$r(G) = \sum_{i=1}^{p} \lambda$

Let F be a distribution functi
f. We define

$\tilde{I}(F) =$

With this definition, (2.3) beco

$r(G)$

where $\tilde{\Phi}$ is the c.d.f. of multi
covariance I, and $*$ denotes con
then $\tilde{I}(F)$ becomes $I(F)$ in [4].
The following lemma gives

LEMMA 2.1. *Let H, F, and H*
and h_n respectively.

1. *If $H(\tilde{x}) = F(\frac{x_1}{m}, \frac{x_2}{m}, \dots$*
2. *Under the assumptions*
   (i) *f, $\frac{\partial f}{\partial x_i}$ are continu*
   (ii) *$\tilde{I}(F) < \infty$,*
   if $H_n \to \delta_o$ (point mass

The parameter space is a s
minimax risk for the case Ω_m
lemmas given next.

Let $\overset{\circ}{W}{}_2^1(\Omega_1)$ be the Sobol
boundary $\partial\Omega_1$ and whose first
by $L \equiv \lambda_1 \frac{\partial^2}{\partial x_1^2} + \dots + \lambda_p \frac{\partial^2}{\partial x_p^2}$ the

From formula (3.3) of [9] with respect to the loss $\sum_{i=1}^{p}\left(\delta_i(\tilde{x})-\theta_i\right)^2$, the Bayes estimator is the posterior mean which is equal to

$$\delta(\tilde{x}, G) = \tilde{x} + \frac{\nabla f_G(\tilde{x})}{f_G(\tilde{x})}.$$

Moreover, $\delta(\tilde{x}, G)$ is also the Bayes estimator with respect to the loss function $\sum_{i=1}^{p} \lambda_i (G_i(\tilde{x}) - \theta_i)^2$ ([3], page 161). Substituting $\delta(\tilde{x}, G)$ into (2.2), the Bayes risk becomes

$$r(G) = \sum_{i=1}^{p} \lambda_i - \int_{\mathbb{R}^p} \sum_{i=1}^{p} \frac{\lambda_i \left(\frac{\partial}{\partial x_i} f_G(\tilde{x})\right)^2}{f_G(\tilde{x})} d\tilde{x}. \tag{2.3}$$

Let F be a distribution function which has an absolutely continuous density f. We define

$$\tilde{I}(F) = \int_{\mathbb{R}^p} \frac{\sum_{i=1}^{p} \lambda_i (\frac{\partial f}{\partial x_i})^2}{f(\tilde{x})} d\tilde{x}.$$

With this definition, (2.3) becomes

$$r(G) = \sum_{i=1}^{p} \lambda_i - \tilde{I}(\tilde{\Phi} * G) \tag{2.4}$$

where $\tilde{\Phi}$ is the c.d.f. of multivariate normal distribution with mean $\tilde{0}$ and covariance I, and $*$ denotes convolution. We note that if $\lambda_i = 1$, $i = 1, 2, \ldots, p$, then $\tilde{I}(F)$ becomes $I(F)$ in [4].

The following lemma gives two properties of $\tilde{I}(F)$.

LEMMA 2.1. *Let H, F, and H_n be distribution functions with densities h, f, and h_n respectively.*

1. *If $H(\tilde{x}) = F(\frac{x_1}{m}, \frac{x_2}{m}, \ldots, \frac{x_p}{m})$ then $\tilde{I}(H) = \frac{1}{m^2}\tilde{I}(F)$.*
2. *Under the assumptions*
   (i) *f, $\frac{\partial f}{\partial x_i}$ are continuous and bounded and h_n is continuous,*
   (ii) *$\tilde{I}(F) < \infty$,*
   if $H_n \to \delta_o$ (point mass at $\tilde{0}$) weakly, then $\tilde{I}(F * H_n) \to \tilde{I}(F)$ as $n \to \infty$.

The parameter space is a sphere of radius m. To get a lower bound of the minimax risk for the case $\Omega_m = \{\tilde{\theta} \in \mathbb{R}^p : \sum_{i=1}^{p} \theta_i^2 \leq m^2\}$ we need the two lemmas given next.

Let $\overset{\circ}{W}{}_2^1(\Omega_1)$ be the Sobolev space of all functions that vanish on the boundary $\partial\Omega_1$ and whose first partial derivatives are in $L_2(\Omega_1)$. We denote by $L \equiv \lambda_1 \frac{\partial^2}{\partial x_1^2} + \ldots + \lambda_p \frac{\partial^2}{\partial x_p^2}$ the partial differential operator.

LEMMA 2.2. *There exists* $f_{1p} \in \overset{\circ}{W}{}_2^1(\Omega_1)$ *such that* $g_{1p} = f_{1p}^2$ *minimizes* $\tilde{I}(G)$ *subject to* $\int_{\Omega_1} g_{1p}(\tilde{x})d\tilde{x} = \int_{\Omega_1} f_{1p}^2(\tilde{x})d\tilde{x} = 1$ *and* $f_{1p}(\tilde{x}) = 0$ *on* $\partial\Omega_1$. *Furthermore,* $\tilde{I}(G_{1p}) = 4\gamma$ *where* γ *is the minimum eigenvalue of the Dirichlet problem in* Ω_1 *for the operator* L,

$$L(u) = -\gamma u \qquad u|_{\partial\Omega_1} = 0.$$

LEMMA 2.3. *Let* G_{1p} *be the distribution on* Ω_1 *with density* $g_{1p}(\theta_1, \theta_2, \dots, \theta_p)$ *and* G_{mp} *be the corresponding distribution scaled up to* Ω_m *with density given by*

$$g_{mp}(\theta_1, \theta_2, \dots, \theta_p) = \frac{1}{m^p}\, g_{1p}\Big(\frac{\theta_1}{m}, \frac{\theta_2}{m}, \dots, \frac{\theta_p}{m}\Big).$$

Then we have

$$r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \frac{4\gamma}{m^2} + o(m^{-2}) \quad as \quad m \to \infty.$$

PROOF. From (2.4) we have

$$r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \tilde{I}(\tilde{\Phi} * G_{mp}).$$

Using property 1 in Lemma 2.1, $r(G_{mp})$ becomes

$$r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \frac{\tilde{I}(\tilde{\Phi}_{1/m} * G_{1p})}{m^2}.$$

Property 2 in Lemma 2.1 together with the fact that $\tilde{I}(G_{mp}) = 4\gamma$ gives

$$r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \frac{4\gamma}{m^2} + o(x^{-2}) \qquad \text{as} \quad m \to \infty. \qquad \square$$

Lemma 2.3 shows that $r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \frac{4\gamma}{m^2} + o(x^{-2})$ is a lower bound of the minimax risk $\rho_p(m)$. This conclusion follows from the fact that every Bayes risk is a lower bound of the minimax risk. $\rho_p(m)$ indeed will equal to $\sum_{i=1}^{p} \lambda_i - \frac{4\,\gamma}{m^2} + o(m^{-2})$ by constructing an estimator whose maximum risk is equal to $r(G_{mp})$.

Suppose, $\{a_m\}$ is a sequence of positive numbers such that $1 > a_m \downarrow 0$, $m^2 a_m^2 \to \infty$, $n = \frac{m}{1-a_m}$ and $v(\tilde{x}) = g_{1p}(\tilde{x}) = f_{1p}^2(\tilde{x})$. Let

$$\psi_i(\tilde{x}) = \frac{-\frac{\partial}{\partial x_i} v(\tilde{x})}{v(\tilde{x})} \qquad \text{if} \quad \|x\| < 1 - a_m^p, \qquad i = 1, 2, \dots, p.$$

We define $\delta_m(\tilde{x}) = \tilde{x} - \frac{1}{n}\left[\psi_1\left(\frac{\tilde{x}}{n}\right), \dots, \psi_p\left(\frac{\tilde{x}}{n}\right)\right]'$.

THEOREM 2.4. *The estimator*

PROOF. We will show that

$$\sup_{\|\theta\| \le m} \Big\{ \Big| R(\tilde{\theta}, \delta_m(\tilde{x})) - \sum_{i=}$$

from which the theorem follow

We recall that if $\|x\| < 1 -$

$$\frac{\partial}{\partial x_i}\psi_i(\tilde{x}) = 2$$

For $\|x\| < 1 - a_m^2$ we have

$$\sum_{i=1}^{p} \lambda_i \psi_i^2(\tilde{x})$$

Now the theorem is equivalent

$$\sup_{\|\theta\| \le m} \Big[\int_{\mathbb{R}^p} \Big\{ \sum_{i=1}^{p} \lambda_i \psi_i^2\Big(\frac{x_1}{n}, \dots$$

$$\cdot \frac{1}{(\sqrt{2\pi})^{p/2}} e^{-\frac{1}{2}\sum_{i=1}^{p}}$$

By virtue of (2.5) we have

$$\sum_{i=1}^{p} \lambda_i \psi_i^2\Big(\frac{x_1}{n}, \dots, \frac{x_p}{n}\Big) - 2\sum_{i=}$$

Hence (2.6) is equivalent to

$$\sup_{\|\theta\| \le m} \Big[\int_{\|x\| > m(1+\alpha_m)}$$

$$= \sup_{\|\theta\| \le m} P \Big[$$

$$= P \Big[|\tilde{x} - \tilde{\theta}|$$

THEOREM 2.4. *The estimator $\delta_m(\tilde{x})$ is second order asymptotically minimax.*

PROOF. We will show that

$$\sup_{\|\theta\|\le m}\left\{\left|R(\tilde{\theta},\delta_m(\tilde{x}))-\sum_{i=1}^{p}\lambda_i+\frac{4\gamma}{m^2}\right|\right\}=o(m^{-2})\qquad \text{as}\quad m\to\infty,$$

from which the theorem follows by lemma 2.3.

We recall that if $\|x\|<1-a_m^2$, then $\psi_i(\tilde{x})=\dfrac{-\frac{\partial v(\tilde{x})}{\partial x_i}}{v(\tilde{x})}$, $i=1,2,\ldots,p$. So

$$\frac{\partial}{\partial x_i}\psi_i(\tilde{x})=2\sum_{i=1}^{p}\lambda_i\frac{\frac{\partial^2 v}{\partial x_i^2}}{v}-\sum_{i=1}^{p}\lambda_i\frac{\left(\frac{\partial v}{\partial x_i}\right)^2}{v^2}.$$

For $\|x\|<1-a_m^2$ we have

$$\sum_{i=1}^{p}\lambda_i\psi_i^2(\tilde{x})-2\sum_{i=1}^{p}\lambda_i\frac{\partial}{\partial x_i}\psi_i(\tilde{x})=-4\gamma. \tag{2.5}$$

Now the theorem is equivalent to

$$\sup_{\|\theta\|\le m}\left[\int_{\mathbb{R}^p}\left\{\sum_{i=1}^{p}\lambda_i\psi_i^2\Big(\frac{x_1}{n},\ldots,\frac{x_p}{n}\Big)-2\sum_{i=1}^{p}n\lambda_i\frac{\partial}{\partial x_i}\psi_i\Big(\frac{x_1}{n},\ldots,\frac{x_p}{n}\Big)+4\gamma\right\}\cdot\frac{1}{(\sqrt{2\pi})^{p/2}}e^{-\frac12\sum_{i=1}^{p}(x_i-\theta_i)^2}d\tilde{x}\right]=o(1)\qquad \text{as } m\to\infty. \tag{2.6}$$

By virtue of (2.5) we have

$$\sum_{i=1}^{p}\lambda_i\psi_i^2\Big(\frac{x_1}{n},\ldots,\frac{x_p}{n}\Big)-2\sum_{i=1}^{p}n\lambda_i\frac{\partial\psi_i}{\partial x_i}\Big(\frac{x_1}{n},\ldots,\frac{x_p}{n}\Big)=\begin{cases}-4\gamma & \text{if } \frac{\|x\|}{n}\le 1-a_m^2\\ 0 & \text{otherwise.}\end{cases}$$

Hence (2.6) is equivalent to

$$\begin{aligned}\sup_{\|\theta\|\le m}&\left[\int_{\|x\|>m(1+\alpha_m)}\frac{1}{(\sqrt{2\pi})^{p/2}}\,e^{-\frac12\left(\sum_{i=1}^{p}(x_i-\theta_i)^2\right)}d\tilde{x}\right]\\ &=\sup_{\|\theta\|\le m}P\left[\sum_{i=1}^{p}x_i^2>m^2(1+a_n)^2\right]\\ &=P\left[|\tilde{x}-\tilde{\theta}|\ge ma_m\right]=o(1)\qquad \text{as}\quad m\to\infty.\end{aligned}$$

□

Case when the parameter space is a cube. Now we assume that

$$\tilde{\theta} \in \Sigma_m = \{\tilde{\theta} : \max_{1 \le i \le p} |\theta_i| \le m\}.$$

LEMMA 2.5. *Let $G_{1p}(\tilde{\theta})$ be the distribution on Σ_1 with density*

$$g_{1p}(\tilde{\theta}) = \cos^2 \frac{\pi}{2}\theta_1 \ \ldots \ \cos^2 \frac{\pi}{2}\theta_p, \quad \tilde{\theta} \in \Sigma_1.$$

Then G_{1p} minimizes $\tilde{I}(F)$ subject to

$$\int_{\Sigma_1} g_{1p}(\tilde{\theta})\, d\tilde{\theta} = 1 \qquad \text{and} \qquad g_{1p}(\tilde{\theta}) = 0 \ \text{ on } \ \partial\Sigma_1.$$

Moreover, $\tilde{I}(G_{1p}) = \pi^2 \sum_{i=1}^{p} \lambda_1$.

LEMMA 2.6. *Let $G_{mp}(\tilde{\theta})$ be the distribution G_{1p} scaled up to Σ_m with density given by*

$$g_{mp}(\tilde{\theta}) = \frac{1}{m^p} \cos^2 \frac{\pi}{2}\frac{\theta_1}{m} \ \ldots \ \cos^2 \frac{\pi}{2}\frac{\theta_p}{m}, \quad \tilde{\theta} \in \Sigma_m.$$

Then we have

$$r(G_{mp}) = \sum_{i=1}^{p} \lambda_i - \frac{\pi^2}{m^2} \sum_{i=1}^{p} \lambda_i + o(m^{-2}).$$

In the next theorem we will construct an estimation whose maximum risk achieves the above lower bound.

THEOREM 2.7. *Let $\{a_m\}$ be a sequence of positive numbers such that*

$$1 > a_m \downarrow 0, \quad m a_m \to \infty, \quad \text{and} \quad n = \frac{m}{(1 - a_m)}.$$

We define

$$\delta_m(\tilde{x}) = \big(\delta_1(\tilde{x}), \ldots, \delta_p(\tilde{x})\big)'$$

where

$$\delta_i(\tilde{x}) = x_i - \frac{\pi}{n} \tan \frac{\pi}{2}\frac{x_i}{n}\, 1_{[-m(1+\alpha_m), m(1+\alpha_m)]}(x_i), \qquad i = 1, 2, \ldots, p.$$

The estimator $\delta_m(\tilde{x})$ is second order asymptotically minimax in the sense that

$$\sup_{\theta \epsilon \Sigma_m} \left\{ \left| R\big(\delta_m(\tilde{x}), \tilde{\theta}\big) = \sum_{i=1}^{p} \lambda_i + \frac{\pi^2}{m^2} \sum_{i=1}^{p} \lambda_i \right| \right\} = o(m^{-2}) \qquad \text{as} \quad m \to \infty.$$

Now we will extend the
observation.
Let x_i be $N(\theta_i, 1)$, $i = 1$
$(\theta_1, \theta_2, \ldots)'$ under the weight

$$L(\tilde{\delta}_d, \tilde{\theta}_d) = \sum_{i=1}^{\infty} \lambda_i (\delta_i$$

It is assumed that

$$\tilde{\theta}_d \in \Sigma_\infty$$

THEOREM 2.8. *Let $\rho_d(m)$ be*

$$\rho_d(m) =$$

Then we have

$$\rho_d(m) = \sum_{i=1}^{\infty} \lambda_i - \frac{\pi}{m}$$

In addition, the estimator $\tilde{\delta}_d$
totically minimax.

PROOF. Step 1: Sup
density given by $g_{mi}(\theta_i) = \frac{1}{m}$
dent. Let $\overline{G}_m = \prod_{i=1}^{\infty} G_{mi}$ be
$\overline{G}_m$ is guaranteed by Ash ([2,
Since (θ_i, x_i), $(x_1, \ldots, x_{i-}$
262], we have

$$E(\theta_i | x_1, \ldots$$

It turns out that

$$E(\theta_i |$$

where$f_{G_{mi}}(x_i) = \phi * g_{mi}(x_i)$.
It can be seen that the E
each i, have Bayes risk equal

$$\lambda_i$$

Now we will extend the results above to a countably infinite dimensional observation.

Let x_i be $N(\theta_i, 1)$, $i = 1, 2, 3, \ldots$. We are interested in estimating $\tilde{\theta}_d = (\theta_1, \theta_2, \ldots)'$ under the weighted quadratic loss function

$$L(\tilde{\delta}_d, \tilde{\theta}_d) = \sum_{i=1}^{\infty} \lambda_i \big(\delta_i(\tilde{x}) - \theta_i\big)^2, \quad \lambda_i > 0 \quad \text{and} \quad \sum_{i=1}^{\infty} \lambda_i < \infty.$$

It is assumed that

$$\tilde{\theta}_d \in \Sigma_\infty = \Big\{\tilde{\theta}_d : \sup_{1 \le i \le \infty} |\theta_i| \le m\Big\}.$$

THEOREM 2.8. *Let $\rho_d(m)$ be the minimax risk in estimating $\tilde{\theta}_d$. So*

$$\rho_d(m) = \inf_{\tilde{\delta}_d \in \mathcal{D}} \sup_{\tilde{\theta}_d \in \Sigma_\infty} E\big(L(\tilde{\delta}_d, \tilde{\theta}_d)\big).$$

Then we have

$$\rho_d(m) = \sum_{i=1}^{\infty} \lambda_i - \frac{\pi^2}{m^2} \sum_{i=1}^{\infty} \lambda_i + o(m^{-2}) \qquad \text{as} \qquad m \to \infty.$$

In addition, the estimator $\tilde{\delta}_d(\tilde{x}) = (\delta_1(\tilde{x}), \delta_2(\tilde{x}), \ldots)'$ is second order asymptotically minimax.

PROOF. Step 1: Suppose $G_{mi}(\theta_i)$ is a distribution on $[-m, m]$ with density given by $g_{mi}(\theta_i) = \frac{1}{m} \cos^2 \frac{\pi}{2} \frac{\theta_i}{m}$, $i = 1, 2, 3, \ldots$, and G_{mi} are independent. Let $\overline{G}_m = \prod_{i=1}^{\infty} G_{mi}$ be the prior distribution on Σ_∞. The existence of $\overline{G}_m$ is guaranteed by Ash ([2, page 231]).

Since (θ_i, x_i), $(x_1, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots)$ are independent using [2, page 262], we have

$$E(\theta_i | x_1, \ldots, x_{i-1}, x_i, x_{i+1} \ldots) = E(\theta_i | x_i).$$

It turns out that

$$E(\theta_i | x_i) = x_i - \frac{\frac{d}{dx_i} f_{G_{mi}}(x_i)}{f_{G_{mi}}(x_i)}$$

where$f_{G_{mi}}(x_i) = \phi * g_{mi}(x_i)$.

It can be seen that the $E(\theta_i | x_i)$ are independent, $i = 1, 2, 3, \ldots$, and for each i, have Bayes risk equal to

$$\lambda_i - \frac{\pi^2}{m^2} \lambda_i + \lambda_i o(m^{-2}).$$

Consequently, with respect to prior $\overline{G}_m$, the Bayes risk is equal to

$$r(\overline{G}_m) = \sum_{i=1}^{\infty} \lambda_i - \frac{\pi^2}{m^2} \sum_{i=1}^{\infty} \lambda_i + o(m^{-2}) \le \rho_d(m). \tag{2.7}$$

By showing that the maximum risk of $\tilde{\delta}_d(\tilde{x})$ asymptotically achieves the lower bound in (2.7), the inequality becomes equality.

Step 2 : Let

$$\psi_i(\tilde{x}) = \frac{\pi}{n} \tan \frac{\pi}{2} \frac{x_i}{n} 1_{[-m(1+\alpha_m), m(1+\alpha_m)]}(x_i) \qquad i = 1, 2, 3, \ldots .$$

Then if follows that

$$R(\delta_m, \tilde{\theta}_d) = \sum_{i=1}^{\infty} \lambda_i + \sum_{i=1}^{\infty} \lambda_i E_{\tilde{\theta}} \Big[\psi_i^2(\tilde{x}) - 2 \frac{\partial}{\partial_{x_i}} \psi_i(\tilde{x}) \Big].$$

From Theorem 2.7 we have

$$\psi_i^2(\tilde{x}) - 2 \frac{\partial}{\partial x_i} \psi_i(\tilde{x}) = -\frac{\pi^2}{n} 1_{[-m(1+\alpha_m), m(1+\alpha_m)]}(x_i) \qquad i = 1, 2, 3, \ldots .$$

Hence the asymptotic minimaxity of $\delta_d(\tilde{x})$ is equivalent to

$$\sum_{i=1}^{\infty} \lambda_i \sup_{-m \le \theta_i \le m} \int_{|x_i| > m(1+\alpha_m)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \theta_i)^2} dx_i = o(1) \text{ as } m \to \infty. \tag{2.8}$$

The left hand side of (2.8) is equal to $\sum_{i=1}^{\infty} \lambda_i [1 - \Phi(ma_m) + \Phi(-ma_m)]$ which tends to 0 as $m \to \infty$. □

3. Application to an arbitrary quadratic loss function

In this section we will present a method to transform the case of an arbitrary quadratic loss function to that of a weighted quadratic loss function so that the problem can be solved by using the results of section 2. This method uses the singular value decomposition of the positive definite matrix to reduce to the case of a diagonal matrix.

Let $\tilde{X}_p$ be $N_p(\tilde{\theta}, I)$ and assume that

$$\tilde{\theta} \in \Omega_m = \Big\{ \tilde{\theta} \in \mathbb{R}^p : \sum_{i=1}^{p} \theta_i^2 \le m^2 \Big\} \text{ or } \tilde{\theta} \in \Sigma_p = \Big\{ \tilde{\theta} \in \mathbb{R}^p : \sup_{1 \le i \le p} |\theta_i| \le m \Big\}$$

It is desired to estimate $\tilde{\theta}$ using an estimator $\delta(\tilde{x})$ under an arbitrary quadratic loss function $L(\tilde{\delta}, \tilde{\theta}) = (\tilde{\delta} - \tilde{\theta})' Q (\tilde{\delta} - \tilde{\theta})$, where Q is a known symmetric positive definite matrix. To transform the above loss function we use the singular value decomposition ([1], p. 84) which is given in the next theorem.

THEOREM 3.1. *Given B posi*
exists a nonsingular matrix F

$F'B$

$F'A$

where $\lambda_1 \ge \ldots \ge \lambda_p$ (≥ 0) *are*
If B is positive definite, t

Using Theorem 3.1 above

$C'Q$

where $\lambda_i > 0$, $i = 1, 2, \ldots, p$,
Then,

$\tilde{Y} \sim N($

Furthermore,

$L(\tilde{\delta}, \tilde{\theta}) = (C\delta(\tilde{y}) - C$

$= (\delta(\tilde{y}) - \theta^*)$

From the above result it
mates based on an arbitrary
case of a weighted quadratic l

[1] Anderson, T. W., *An Introduc*
Wiley, New York, 1984.
[2] Ash, R. B., *Real Analysis and*
[3] Berger, J. O., *Statistical Decis*
[4] Bickel, P. J., *Minimax estima*
rameter space is restricted, An
[5] Brown, L. D. and Gajek, L., *I*
18(1990), 1578–1594.
[6] Clevenson, M. L. and Zidek, J
Poisson laws, JASA, **70**(1975)
[7] Gilbarg, D. and Trudinger, N
Order, Springer-Verlag, New Y
[8] Johnstone, I. M. and Mac Gibb
vector, Ann. Statist., **20**(1992)
[9] Stein, C., *Estimation of the m*
9(1991), 1135–1151.
[10] Subanar et al., *Minimax Estim*
nical Report, Faculty of Scienc

yes risk is equal to

$m^{-2}) \leq \rho_d(m).$ (2.7)

mptotically achieves the lower

$x_i) \qquad i = 1, 2, 3, \ldots .$

$\tilde{x}) - 2\frac{\partial}{\partial_{x_i}}\psi_i(\tilde{x})\Big].$

$_{m})](x_i) \qquad i = 1, 2, 3, \ldots .$

uivalent to

$dx_i = o(1)$ as $m \to \infty$. (2.8)

$- \Phi(ma_m) + \Phi(-ma_m)]$ which □

dratic loss function

transform the case of an arbi-
ted quadratic loss function so
ults of section 2. This method
itive definite matrix to reduce

$= \Big\{\tilde{\theta} \in \mathbb{R}^p : \sup_{1\leq i\leq p} |\theta_i| \leq m\Big\}$

) under an arbitrary quadratic
is a known symmetric positive
ction we use the singular value
next theorem.

THEOREM 3.1. *Given B positive semi definite and A positive definite, there exists a nonsingular matrix F such that*

$$F'BF = Diag(\lambda_1, \ldots, \lambda_p) \tag{3.1}$$

$$F'AF = I \tag{3.2}$$

where $\lambda_1 \geq \ldots \geq \lambda_p$ (≥ 0) are the roots of $|B - \lambda A| = 0$.

If B is positive definite, then $\lambda_i > 0$, $i = 1, 2, \ldots, p$.

Using Theorem 3.1 above, there exists a non singular matrix C such that

$$C'QC = D(\lambda_1, \lambda_2, \ldots, \lambda_p)$$

where $\lambda_i > 0$, $i = 1, 2, \ldots, p$, and $C'C = I$. Let $\tilde{X} = C\tilde{Y}$ and $\delta(\tilde{x}) = C\delta(\tilde{y})$. Then,

$$\tilde{Y} \sim N(\tilde{\theta}^*, I) \quad \text{with} \quad \tilde{\theta}^* = C^{-1}\tilde{\theta}.$$

Furthermore,

$$L(\tilde{\delta}, \tilde{\theta}) = \big(C\delta(\tilde{y}) - C\tilde{\theta}^*\big)'Q\big(C\delta(\tilde{y}) - C\tilde{\theta}^*\big)$$

$$= \big(\delta(\tilde{y}) - \theta^*)'C'QC(\delta(\tilde{y}) - \theta^*\big) = \sum_{i=1}^{p} \lambda_i \big(\delta(\tilde{y}) - \theta^*\big)^2$$

From the above result it can be seen that construction of minimax estimates based on an arbitrary quadratic loss function could be reduced to the case of a weighted quadratic loss function.

References

[1] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, 2nd ed., John Wiley, New York, 1984.

[2] Ash, R. B., *Real Analysis and Probability*, Academic Press, New York, 1972.

[3] Berger, J. O., *Statistical Decision Theory*, Springer-Verlag, New York, 1980.

[4] Bickel, P. J., *Minimax estimation of the mean of a normal distribution when the parameter space is restricted*, Ann. Statist., **9**(1981), 1301–1309.

[5] Brown, L. D. and Gajek, L., *Information inequalities for the Bayes risk*, Ann. Statist., **18**(1990), 1578–1594.

[6] Clevenson, M. L. and Zidek, J., *Simultaneous estimation of the mean of independent Poisson laws*, JASA, **70**(1975), 648–705.

[7] Giilbarg, D. and Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, New York, 1977.

[8] Johnstone, I. M. and Mac Gibbon, K. B., *Minimax Estimation of a constrained Poisson vector*, Ann. Statist., **20**(1992), 807–831.

[9] Stein, C., *Estimation of the mean of a multivariate normal distribution*. Ann. Statist., **9**(1991), 1135–1151.

[10] Subanar et al., *Minimax Estimation with respect to the symmetric loss function*, Technical Report, Faculty of Science & Math, Gadjah Mada University, 1993.