On the Minimax Estimator of a Bounded Normal Mean

Éric Marchand[*,†] François Perron[‡,†]

CRM-2756

September 2001

[*]Department of Mathematics and Statistics, University of New Brunswick, P.0. Box 4400, Fredericton, N.B. CANADA, E3B 5A3

[†]Research supported by NSERC of Canada

[‡]Département de mathématiques et de statistique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Qc, CANADA H3C 3J7 perronf@dms.umontreal.ca

Abstract

For estimating under squared-error loss the mean of a p-variate normal distribution when this mean lies in a ball of radius m centered at the origin and the covariance matrix is equal to the identity matrix, it is shown that the Bayes estimator with respect to a uniformly distributed prior on the boundary of the parameter space (δ_{BU}) is minimax whenever $m \leq \sqrt{p}$. Further descriptions of the cutoff points of small enough radiuses (i.e., $m \leq m_0(p)$) for δ_{BU} to be minimax are given. These include lower bounds and the large dimension p limiting behaviour of $m_0(p)/\sqrt{p}$. Finally, implications for the associated minimax risk are described.

AMS 2000 subject classifications. 62F10, 62F15, 62F30

Keywords and phrases: Minimax estimator, restricted parameter space, multivariate normal distribution, squared error loss

1 Introduction

Consider the problem of estimating, based on one observation X, the mean θ of $X \sim N_p(\theta, I_p)$ under squared error loss, and with the constrained parameter space $\Theta(m) = \{\theta \in \mathbb{R}^p : \|\theta\| \leq m\}$ for some m fixed, $m > 0$. Marchand and Perron (2001) show that the Bayes estimator δ_{BU} with respect to the boundary uniform prior on $\partial\Theta(m)$ dominates the maximum likelihood estimator $\delta_{\text{mle}}(x) = \left(\frac{m}{\|x\|} \wedge 1\right) x$ whenever $m \leq \sqrt{p}$. On the other hand, it follows from the work of Das Gupta (1985) that δ_{BU} is minimax for small enough m, say $m \leq m_0(p)$. An interesting question is whether the condition $m \leq \sqrt{p}$ is also sufficient for the minimaxity of δ_{BU}; in other words whether or not $m_0(p) \geq \sqrt{p}$.

For $p = 1$, the answer is available from Casella and Strawderman (1981) who report that $m_0(1) \approx 1.05674$. (In fact, they showed as well that δ_{BU} dominates δ_{mle} for $p = 1$ and $m \leq \sqrt{1}$). Berry (1990) gives $m_0(2) = 1.09318\sqrt{2}$ and $m_0(3) = 1.1015783\sqrt{3}$ which implies that for p=2,3, the condition $m \leq \sqrt{p}$ is also sufficient for the minimaxity of δ_{BU}. Although, for fixed p, one can assess numerically whether the inequality $m_0(p) \geq \sqrt{p}$ holds, further analytical specifications of $m_0(p)$ are unavailable, and the minimaxity of δ_{BU} for $m \leq \sqrt{p}; p \geq 4$; remains an open question.

In this paper, we show that indeed $m_0(p) \geq \sqrt{p}$ for all $p \geq 4$. As further corollaries, (i) we give implicit and explicit lower bounds for $m_0(p)$, (ii) establish that $\lim_{p\to\infty} m_0(p)/\sqrt{p} = \kappa \approx 1.150963925$, and (iii) discuss implications for the minimax risk whenever δ_{BU} is minimax, in particular when the dimension p is large.

The occurrence of the simple and common cutoff point $\sqrt{p}$ for the parameter space $\Theta(m)$ to have a small enough radius m in order that δ_{BU} both (i) dominate δ_{mle} and (ii) be minimax, is of intrinsic interest. Interestingly, it also represents the cutoff point of radiuses m (i.e., $m \leq \sqrt{p}$) such that all estimators δ of θ that take values in the parameter space $\Theta(m)$ dominate the unbiased estimator, since $R(\theta, \delta) \leq m^2$ and since the risk of the unbiased estimator is p (in fact, we only require $X - \theta$ to be spherically symmetric for this to be true).

2 Preliminaries and Lower Bounds for $m_0(p)$

As in Marchand and Perron (2001),

$$\delta_{BU}(x) = \overline{g}_m(\|x\|)\frac{x}{\|x\|} \quad ; \tag{1}$$

where for $r > 0,\ \lambda \geq 0$,

$$\overline{g}_\lambda(r) = E_\theta\left[\frac{\theta' X}{\|X\|}\,\middle|\, \|X\| = r\right], \text{ with } \|\theta\| = \lambda. \tag{2}$$

It is known (e.g., Berry 1990, Robert 1990) that

$$\overline{g}_\lambda(r) = \lambda \rho_{\frac{p}{2}-1}(\lambda r), \tag{3}$$

where $\rho_\nu(t) = \dfrac{I_{\nu+1}(t)}{I_\nu(t)}$; $\nu > -1$; I_ν representing the modified Bessel function of order ν. Moreover, as in Marchand and Perron (2001), we will exploit the following properties of $\rho_{\frac{p}{2}-1}$.

Lemma 1. (Watson, 1982) For all $p \geq 1, \rho_{\frac{p}{2}-1}(t)$ is increasing and concave in t; $\rho_{\frac{p}{2}-1}(t)/t$ is decreasing in t; $\rho_{\frac{p}{2}-1}(0) = 0$ and $\rho_{\frac{p}{2}-1}(t) \to 1$ as $t \to \infty$.

Recalling a familiar criterion for minimaxity, it follows that δ_{BU} will be minimax if the supremum of its risk function, $R(\theta, \delta_{BU})$, is attained on the support of the associated prior, that is to say, the boundary $\partial\Theta(m)$. Moreover, Berry (1990) established that $\sup_{\theta\in\Theta(m)}\{R(\theta, \delta_{BU})\}$ is attained at either the origin ($\theta = 0$) or the boundary ($\|\theta\| = m$). Hence, our starting point.

Lemma 2. (Berry, 1990) A necessary and sufficient condition for δ_{BU} to be minimax is:

$$R(0,\delta_{BU}) \leq R(\theta,\delta_{BU}) \ , \ \text{ with } \ \|\theta\| = m. \tag{4}$$

In the following we will denote for convenience $R(0,\delta_{BU})$ and $R(\theta,\delta_{BU})$ with $\|\theta\| = m$, as $R_1(m,p)$ and $R_2(m,p)$ respectively.

Lemma 3. Condition (4) for δ_{BU} to be minimax is equivalent to

$$E_0\left[\rho^2_{\frac{p}{2}-1}\left(m\sqrt{R^2}\right)\right] + E_{m^2}\left[\rho^2_{\frac{p}{2}-1}\left(m\sqrt{R^2}\right)\right] \leq 1 \ ; \tag{5}$$

where the expectation E_γ is taken with respect to $R^2 \sim \chi^2_p(\gamma)$.

Proof. From expression (1) we have $\delta_{BU}(x) = \bar{g}_m(\|x\|)\ \frac{x}{\|x\|}$. We may write

$$\begin{aligned} R(\theta,\delta_{BU}) &= E_\theta[\|\delta_{BU}(X) - \theta\|^2] \\ &= \theta'\theta + E_\theta[\bar{g}^2_m(\|X\|)] - 2E_\theta\left[\theta' X \ \frac{\bar{g}_m(\|X\|)}{\|X\|}\right] \\ &= \theta'\theta + E_{\theta'\theta}[\bar{g}^2_m(\sqrt{R^2})] - 2E_{\theta'\theta}[\bar{g}_m(\sqrt{R^2})\bar{g}_\lambda(\sqrt{R^2})] \end{aligned}$$

given (2). From this and (3), we have

$$R_1(m,p) = E_0[m^2\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] \ , \ \text{ and } \ R_2(m,p) = m^2 - E_{m^2}[m^2\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})]. \tag{6}$$

The result now follows directly from (4) and (6).

Given the properties of $\rho_{\frac{p}{2}-1}$ of Lemma 1 and the monotone likelihood ratio of the density of R^2, with m viewed as the parameter it is easy to see that the left hand side of the inequality in expression (5) is increasing in m. Therefore, the inequality in expression (5) will be satisfied if and only if $m \leq m_0(p)$ for some value of $m_0(p)$.

Below, we establish in Lemma 5 simple upper bounds for the quantities $E_\gamma[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})]$; $\gamma = 0, m^2$; which will lead to sufficient conditions (Corollary 6) for the minimaxity of δ_{BU}; at least one of which is shown to be verified in Theorem 7 for $p \geq 4$. The results below exploit the following bounds for $\rho^2_\nu(x)$, which we express in terms of the function $L(a,b) = \{\frac{a}{2} + \sqrt{1 + (\frac{b}{2})^2}\}^{-2}$.

Lemma 4. (Amos, 1974, equation 11) For $\nu \geq 0$,

$$L(\frac{2(\nu+1)}{x}, \frac{2(\nu+1)}{x}) \leq \rho^2_\nu(x) \leq L(\frac{2\nu}{x}, \frac{2(\nu+2)}{x}).$$

Lemma 5.

(a) $E_\gamma[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] \leq \rho^2_{\frac{p}{2}-1}(m\sqrt{\rho+\gamma})$;

(b) For $m = c\sqrt{p}, \ p > 1$,

$$E_0[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] \leq \rho^2_{\frac{p}{2}-1}(cp) \leq L(\frac{1-2/p}{c}, \frac{1+2/p}{c}),$$

and

$$E_{m^2}[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] \leq \rho^2_{\frac{p}{2}-1}(pc\sqrt{1+c^2}) \leq L(\frac{1-2/p}{c\sqrt{1+c^2}}, \frac{1+2/p}{c\sqrt{1+c^2}});$$

(c) For $m = c\sqrt{p},\ p > 1$,

$$E_0[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] + E_{m^2}[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] \leq 2\rho^2_{\frac{p}{2}-1}(pc\sqrt{1+\frac{c^2}{2}}) \leq 2L(\frac{1-2/p}{c\sqrt{1+\frac{c^2}{2}}}, \frac{1+2/p}{c\sqrt{1+\frac{c^2}{2}}}).$$

Proof. Part (a) follows from the concavity of $\rho^2_{\frac{p}{2}-1}(\sqrt{y})$ in y, Jensen's inequality, and the property $E_\gamma[R^2] = p + \gamma$. The concavity follows by the direct evaluation

$$\frac{\partial}{\partial y}\{\rho^2_{\frac{p}{2}-1}(\sqrt{y})\} = \left[\frac{\rho_{\frac{p}{2}-1}(\sqrt{y})}{\sqrt{y}}\right]\left[\frac{\partial}{\partial y}\{\rho_{\frac{p}{2}-1}(\sqrt{y})\}\right],$$

which, by virtue of Lemma 1, is the product of two decreasing (and positive) functions of y. Part (b) is derived from Lemma 4, and a specification of (a) for the radius $m = c\sqrt{p}$. Finally, part (c) follows from: (i) part (b); (ii) the concavity of $\rho^2_{\frac{p}{2}-1}(\sqrt{y})$ in y which implies $\frac{1}{2}\rho^2_{\frac{p}{2}-1}(\sqrt{p^2c^2}) + \frac{1}{2}\rho^2_{\frac{p}{2}-1}(\sqrt{p^2c^2(1+c^2)}) \leq \rho^2_{\frac{p}{2}-1}(\sqrt{p^2c^2(1+\frac{c^2}{2})}\ )$; and (iii) Lemma 4.

Corollary 6. For $m = c\sqrt{p},\ p > 1$, each one of the following three conditions is sufficient for δ_{BU} to be minimax:

(a)

$$\rho^2_{\frac{p}{2}-1}(pc) + \rho^2_{\frac{p}{2}-1}(pc\sqrt{1+c^2}) \leq 1\ ; \tag{7}$$

(b)

$$L(\frac{1-2/p}{c}, \frac{1+2/p}{c}) + L(\frac{1-2/p}{c\sqrt{1+c^2}}, \frac{1+2/p}{c\sqrt{1+c^2}}) \leq 1\ ; \tag{8}$$

(c)

$$c \leq c_3(p) = \sqrt{\sqrt{1+d^2(p)}\ - 1} \tag{9}$$

with $d(p) = (1-\frac{2}{p}) + \sqrt{1+(\frac{2}{p})^2}$.

Proof. Parts (a) and (b) follow immediately from part (b) of Lemma 5, while part (c) comes from part (c) of Lemma 5. For the derivation of part (c) let $\mu = c\sqrt{1+c^2/2}$. We shall have

$$L(\frac{1-2/p}{\mu}, \frac{1+2/p}{\mu}) \leq 1/2$$

for $\mu > 0$ if and only if

$$0 < \mu \leq d(p)/\sqrt{2}.$$

Finally, if we replace μ by its value in terms of c then we obtain our result.

Theorem 7. $m_0(p) \geq \sqrt{p}$ for all $p \geq 1$.

Proof. The cases $p = 1, 2, 3$ were dealt with in the Introduction. For $p = 4$ and $c = 1$, the left-hand side of (7) is $\rho_1^2(4) + \rho_1^2(4\sqrt{2}) = 0.99459$, whence the result for $p = 4$. Similarly, $\rho_1^2(5) + \rho_1^2(5\sqrt{2}) = 0.97058$ and $\rho_1^2(6) + \rho_1^2(6\sqrt{2}) = 0.95495$, whence the result for $p = 5$ and $p = 6$. Finally, we have $c_3(7) = 1.0096039$, and the proof is completed by observing that $c_3(p)$ is increasing in p ; $p \geq 1$; which implies that $c_3(p) \geq c_3(7) > 1$, for all $p \geq 7$.

The sufficient conditions of Corollary 6 yield lower bounds for $m_0(p)$. Indeed, letting $c_0(p) = m_0(p)/\sqrt{p}$, we have $c_0(p) \geq c_1(p) \geq c_2(p) \geq c_3(p)$, where $c_1(p)$ and $c_2(p)$ are the positive solutions in c of the equations $\rho^2_{\frac{p}{2}-1}(pc) + \rho^2_{\frac{p}{2}-1}(pc\sqrt{1+c^2}) = 1$ and $L(\frac{1-2/p}{c}, \frac{1+2/p}{c}) + L(\frac{1-2/p}{c\sqrt{1+c^2}}, \frac{1+2/p}{c\sqrt{1+c^2}}) = 1$ respectively, and $c_3(p)$ is given in (9). Although $c_1(p)$ and $c_2(p)$ are not available in closed forms, they will be of use below in

Section 3 where they will match the limiting behaviour in p of $c_0(p)$. The bound $c_3(p)$ has the advantage of being explicit, but is not as efficient as the other two. Moreover, $\lim_{p\to\infty} c_3(p) = \sqrt{\sqrt{5}-1} = 1.1117859$ which fails to match the corresponding limit for $c_0(p)$ (see Corollary 9). Figure 1 contrasts these lower bounds with $c_0(p)$, and suggests that the lower bounds, as well as the limit in Corollary 9, yield reasonable approximations for $c_0(p)$. Note that the expression $L((1-2/p)/c, (1+2/p)/c)$ is decreasing in p for all $c > 0$ and it is increasing in c for all $p > 1$. Therefore, $c_2(p)$ is increasing in p for all $p > 1$. We conjecture that both $c_1(p)$ and $c_0(p)$ are increasing in p as well.

3 Limiting Behaviour of $m_0(p)$ and Implications for the Minimax Risk

In this section, we study the large dimension p behaviour of $m_0(p)$. Corollary 9 gives the structure of $m_0(p)$ for large p. It is followed by a discussion on the minimax risk. Before proceeding, we will require the following intermediate result.

Lemma 8. If $m = c\sqrt{p}$ then for any fixed $d > 0$, $\lim_{p\to\infty} E_{d^2p}[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] = L(\frac{1}{c\sqrt{1+d^2}}, \frac{1}{c\sqrt{1+d^2}})$.

Proof. By virtue of Lemma 4, if $p > 1$ and $\delta \in [2/p, 1]$, then

$$L(\frac{1}{cy}, \frac{1}{cy}) \le \rho^2_{\frac{p}{2}-1}(m\sqrt{R^2}) \le L(\frac{1-\frac{2}{p}}{cy}, \frac{1+\frac{2}{p}}{cy}) < L(\frac{1-\delta}{cy}, \frac{1}{cy}) \ , \tag{10}$$

with $y = \sqrt{\frac{R^2}{p}}$. Given that $\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})$ is bounded by two continuous and bounded functions of y, and that $Y =^d \sqrt{\frac{R^2}{p}}$ converges in probability to $\sqrt{1+\frac{\gamma}{p}} = \sqrt{1+d^2}$ as p tends to infinity considering that $\|\theta\|^2/p = d^2$ and d^2 is fixed, the desired result holds with δ being made arbitrarily small.

Corollary 9. $\lim_{p\to\infty} \frac{m_0(p)}{\sqrt{p}} = \kappa \approx 1.150963925,$ [4] where κ^2 is the real root of the polynomial $s^3 - s - 1$.

Proof. Let $G(c) = \lim_{p\to\infty}\{E_0[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})] + E_{c^2p}[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})]\}$. Lemma 8 tells us that

$$G(c) = L(\frac{1}{c}, \frac{1}{c}) + L(\frac{1}{c\sqrt{1+c^2}}, \frac{1}{c\sqrt{1+c^2}}) \ ;$$

and since the function G is continuous and increasing in c, with $\lim_{c\to 0} G(c) = 0$ and $\lim_{c\to\infty} G(c) = 2$, there is a unique positive solution to the equation $G(c) = 1$, which is κ. By using the fact that

$$L(\frac{1}{c\sqrt{1+c^2}}, \frac{1}{c\sqrt{1+c^2}}) = \frac{c^2}{c^2+1}, \tag{11}$$

we obtain after a few manipulations that κ has to be a root of $c^6 - c^2 - 1$, yielding the result.

Remark 10 (Implications for the minimax risk)

Whenever δ_{BU} is minimax, the minimax risk is given in (6) by $R_2(m,p)$. Using the upper and lower bounds for $E_{m^2}[\rho^2_{\frac{p}{2}-1}(m\sqrt{R^2})]$, which are available from (10) and part (b) of Lemma 5 respectively, we have for $m = c\sqrt{p}$ and $Y =^d \sqrt{\frac{R^2}{p}}$

$$1 - L(\frac{1-2/p}{c\sqrt{1+c^2}}, \frac{1+2/p}{c\sqrt{1+c^2}}) \le \frac{R_2(m,p)}{m^2} \le 1 - E_{m^2}[L(\frac{1}{cY}, \frac{1}{cY})], \tag{12}$$

which represent bounds on the minimax risk whenever $m \le m_0(p)$. Moreover, the following identity may be inferred from Lemma 8, expressions (6) and (11):

$$\lim_{p\to\infty} \frac{R_2(c\sqrt{p}, p)}{c^2 p} = \frac{1}{1+c^2}. \tag{13}$$

Finally, we point out that Marchand (1993) gives $R_2(m,p) \ge \frac{(p-1)m^2}{p+m^2}$, which yields an alternative lower bound to the one in (12), namely $\frac{p-1}{p}\frac{1}{1+c^2}$, and which matches the lower limit (13) $\frac{1}{1+c^2}$ when $p \to \infty$.

[4] More precisely: $\kappa^2 = a^{1/3} + (1-a)^{1/3}$ where $a = 1/2 - \sqrt{23/108}$.

Acknowledgements

The authors are very grateful to the referee for his comments and suggestions.

References

Amos, D.E. (1974). Computation of Modified Bessel functions and their ratios. *Mathematics of Computation*, **28**, 239-251.

Berry, C. (1990). Minimax estimation of a bounded normal mean vector. *Journal of Multivariate Analysis*, **35**, 130-139.

Casella, G. & Strawderman, W.E. (1981). Estimating a bounded normal mean. *Annals of Statistics*, **9**, 870-878.

DasGupta, A. (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set. *Sankhyā Ser. A*. **47**, 326-332.

Marchand, É. & Perron, F. (2001). Improving on the MLE of a bounded normal mean. *Annals of Statistics*, to appear.

Marchand, É. (1993). Estimation of a multivariate mean with constraints on the norm. *Canadian Journal of Statistics*, **21**, 359-366.

Robert, C. (1990). Modified Bessel functions and their applications in probability and statistics. *Statistics and Probatility Letters*, **9**, 155-161.

Watson, G.S. (1983). Statistics on Spheres. *Wiley*, New York.

Figure 1: Graph of $c_0(p) = \frac{m_0(p)}{\sqrt{p}}$ and various lower bounds.