

Asymptotically Minimax Estimation of Order-Constrained Parameters and Eigenfunctions of the Laplacian on the Ball

Adam Korányi

*Lehman College
Department of Mathematics
Bronx, New York 10468, U.S.A.
adam@alpha.lehman.cuny.edu*

K. Brenda MacGibbon

*GERAD and Université du Québec à Montréal
Département de mathématiques
C.P. 8888, Succursale Centre-ville
Montréal, Canada H3C 3P8
macgibbon.brenda@uqam.ca*

December, 1999
Revised: May, 2001

Les Cahiers du GERAD

G-99-50

Copyright © 2001 GERAD

Abstract

Inspired by previous results on asymptotic minimax estimation for a ball of increasing radius in $\mathbb{R}^n$, we study the analogous problem for domains of importance in order-restricted inference. In particular, we study domains that are formed by the intersection of a ball and a fundamental chamber of a finite reflection group in $\mathbb{R}^n$. We show (1) how to obtain the principal eigenfunction of such a domain and asymptotically, the related least favourable distribution for the associated minimax problem, (2) that the order and positivity constraints in the usual statistical problems generate such chambers and (3) and in an analogous way to the work of Bickel (1981) in the one dimensional case, how to find the asymptotic minimax risk and the second order asymptotic minimax estimate for such a domain.

Résumé

Inspiré par les résultats existants sur l'estimation asymptotique minimax pour une boule de rayon croissant dans $\mathbb{R}^n$, nous étudions le problème analogue pour certains domaines importants en inférence avec des contraintes d'ordre. Plus particulièrement, nous étudions les domaines formés par l'intersection d'une boule et d'une chambre fondamentale d'un groupe de réflexions dans $\mathbb{R}^n$. Nous montrons (1) comment obtenir la fonction propre principale d'un tel domaine et la loi *a priori* la moins favorable asymptotiquement pour le problème du minimax associé, (2) que les contraintes d'ordre et de positivité dans les problèmes statistiques usuels engendrent de telles chambres et (3) d'une manière analogue au travail de Bickel (1981) dans le cas unidimensionnel, comment trouver le risque asymptotiquement minimax et l'estimateur asymptotiquement minimax du second ordre sur un tel domaine.

1 Introduction

In many parametric statistical estimation problems, there is definite prior information concerning the values of a parameter vector (ω). There may be bounds on the individual components (ω)_{*i*} such as “the norm of ω is at most c ” or “a subset of the ω_i are non-negative” or the “ ω_i ’s are non-increasing in i ”. Many computationally feasible estimation methods have been developed to capitalize on such information. How could one theoretically compare the performance of various possible estimators when such prior information is present? One common, admittedly conservative, approach is the worst-case analysis: given some error measure, compute the maximum expected error over the restricted parameter space, and then seek the estimator that minimizes this maximum risk. The resulting best or minimax risk provides a benchmark against which to measure other estimators.

Here we recall various definitions of statistical decision theory associated with the estimation of an unknown parameter vector ω given an observation of $x \in \mathbf{X}$, where $\mathbf{X}$ is a random vector whose distribution depends on the parameter $\omega \in \Omega$, the parameter space. A solution consists of a nonrandomized estimator or decision procedure δ , which is a measurable function from the sample space $\mathbf{X} \rightarrow \Omega$. Let $\mathcal{A}$ denote the space of all possible estimates. A risk function $R(\delta, \omega)$ characterizes the performance of a decision procedure δ for each value of the parameter ω . The risk function is usually defined in terms of an underlying loss function $L(\delta, \omega)$ which maps $\mathcal{A} \times \Omega \rightarrow \mathbb{R}^+ \cup \{0\}$ (where $\mathbb{R}^+$ is the positive real line). To be able to confine our attention to nonrandomized estimators, it will be assumed that the loss function is convex in δ and that $\mathcal{A}$ and Ω are also convex. The loss function will usually be assumed to equal the quadratic loss function $\sum (\delta_i - \omega_i)^2$. The risk of an estimator δ when ω is true, $R(\delta, \omega)$, is then the average loss incurred from using δ ; that is, $R(\delta, \omega) = E_\omega L(\delta(X), \omega)$. An estimator δ^* is minimax for the above problem, if $\sup_{\omega \in \Omega} R(\delta^*, \omega) \leq \sup_{\omega \in \Omega} R(\delta, \omega)$ for all $\delta \in \mathcal{A}$. We will let $\rho(\Omega)$ denote the minimax risk on Ω ; i.e., $\rho(\Omega) = \inf_{\delta \in \mathcal{A}} \sup_{\omega \in \Omega} R(\omega, \delta)$. Note that if $\Omega = \mathbb{R}^n$ then for the normal mean estimation problem $\rho(\mathbb{R}^n) = n$.

Minimax problems are often solved by considering the corresponding Bayes problems. A distribution or prior probability measure π is specified on the parameter space Ω , and a measure of the performance of a procedure δ is given by its Bayes risk

$$r(\delta, \pi) = \int_{\Omega} R(\delta, \omega) \pi(d\omega).$$

Now δ_π is called the Bayes procedure with respect to the prior probability measure π if δ_π minimizes the Bayes risk. The Bayes risk $r(\pi)$ of a distribution or prior probability measure π on Ω is defined as $r(\pi) = r(\delta_\pi, \pi)$. A distribution or prior probability measure π^* is “least favorable” if its Bayes risk is greater than or equal to that of any other distribution. Subject to the decision problem satisfying sufficient regularity conditions, a least favorable prior distribution exists and the corresponding Bayes procedure is minimax (see Wald (1950), Ferguson (1967), Brown (1979), and Kempthorne (1981)).

Here we consider the normal mean estimation problem; that is, X is a $N_n(\omega, I_n)$ random vector where n represents the dimension of the space, I_n the $(n \times n)$ identity matrix and let us assume that the vector ω is contained in the parameter space Ω which is a closed bounded convex domain in $\mathbb{R}^n$.

Casella and Strawderman (1981) first found exact minimax estimates for intervals of the form $[-t, t]$ for small t for this problem in one dimension. (See also Zinzius (1981).) Donoho, Liu and MacGibbon (1990) studied minimax estimates and affine-minimax estimates for the problem of estimating the mean of a standard Gaussian shift when the mean is known to lie in an orthosymmetric, convex and quadratically convex set in ℓ_2 . However, for many closed bounded domains Ω of $\mathbb{R}^n$ (satisfying $\Omega = \overline{int\Omega}$), the finding of exact minimax solutions is analytically intractable. A more soluble analytical problem involves the study of asymptotically minimax estimates; that is, we can consider the asymptotic behaviour of the minimax risk $\rho(t\Omega) = \inf_{\delta} \sup_{\omega \in t\Omega} E_{\omega} L(\delta(X), \omega)$ for positive t and the construction of such asymptotically minimax estimators as $t \rightarrow \infty$.

The connection between asymptotic minimax risk and the principal eigenvalue of an elliptic equation was first elaborated in a series of papers by Levit (1980, 1982, 1985a) and Berkin and Levit (1980) which studied asymptotic second-order minimax estimators under a general class of loss functions in Gaussian and locally asymptotic Gaussian settings, and connected this problem with the principal eigenvalue of the Laplace (or more generally, second order uniformly elliptic) equation in the domain in which the parameter lies. Bickel (1981) independently derived the results for intervals and balls in the Gaussian setting for squared error loss, obtaining explicit second-order asymptotically minimax estimates by suitably rescaling eigenfunctions of the sphere. Johnstone and MacGibbon (1992, 1993) related the problem of finding asymptotic minimax estimates of a bounded Poisson vector to the Gaussian one.

The solution to the Dirichlet problem is easily seen to be related to asymptotically minimax estimation of a normal mean vector (see, e.g. Berkin and Levit (1980), Levit (1980, 1982, 1985a), Johnstone and MacGibbon (1993)). Essentially, finding a minimax estimator with respect to a density f on Ω is equivalent to minimizing the Fisher information $I(f) = \int f^{-1} |\nabla f|^2$. Since, Fisher information I and the energy functional I^* in the classical Dirichlet problem are related by $I^*(v) = \frac{I(v^2)}{4}$ the problem of minimax estimation becomes a search for principal eigenfunctions of the laplacian.

More precisely, using the solution to the classical Dirichlet problem for a bounded domain (not too irregular) Ω in $\mathbb{R}^n$ (see, e.g., Gilbarg and Trudinger (1983)), we have the following general theorem (for a discussion of the proof and applications, see e.g., Levit and Berkin (1980), Johnstone and MacGibbon (1992, 1993)).

Theorem 1 For Ω as above and t positive,

(i) The minimax risk for the n -dimensional problem just described is

$$\rho(t\Omega) = \inf_{\delta \in D} \sup_{\omega \in t\Omega} E_{\omega} \sum_{i=1}^n [\delta_i(X) - \omega_i]^2 = n - 4t^{-2}\lambda(\Omega) + o(t^{-2}), \quad (1)$$

where $\lambda(\Omega)$ denotes the minimum eigenvalue of the Laplace operator $\Delta = \sum_{j=1}^n \partial^2 / \partial \omega_j^2$ on Ω , i.e. the smallest λ for which the equation

$$\begin{aligned} \Delta v(\omega) &= -\lambda v(\omega) & \omega &\in \text{int}\Omega \\ v &= 0 & \omega &\in \partial\Omega \end{aligned} \quad (2)$$

has a non-zero solution. The eigenspace corresponding to $\lambda(\Omega)$ is one-dimensional, and the corresponding eigenfunction $v_{\Omega}(\omega) = v(\omega, \Omega)$ (or $-v_{\Omega}(\omega)$) is strictly positive on Ω .

Assume that v_{Ω} is normalised so that $\int_{\Omega} v_{\Omega}^2 = 1$.

(ii) Let $G_t^*(d\omega)$ denote a least favorable prior distribution for the region $\Omega_t = t\Omega$ ($t > 0$) and $G_{1,\Omega}^{*t}(d\omega)$ the corresponding prior rescaled to Ω . If we let $g_1(\omega) = v_{\Omega}^2(\omega)$ for $\omega \in \Omega$ and $g_1(\omega) = 0$ for $\omega \notin \Omega$, then the probability measure $G_1(d\omega) = g_1(\omega)d\omega$ is the weak limit of the (rescaled) least favorable distributions $G_{1,\Omega}^{*t}$. If $G_t(d\omega)$ represents the distribution corresponding to $G_1(d\omega)$ scaled up to $t\Omega$, that is, whose density is given by $g_t(x) = t^{-n}g_{\Omega}(xt^{-1})$ where $g_{\Omega}(\omega) = g_1(\omega)$, then the G_t are approximately least favorable in the following sense:

$$\rho(t) = r(G_t) + o(t^{-2}).$$

The proof depends on an easily proved identity, called Brown's identity, (see Brown (1979), Bickel (1981) Brown and Gajek (1990)) which says that for any prior distribution G one has

$$\delta_G(x) = x + \text{grad} \log (g * \phi_n) \quad (3)$$

where g is the density of G (which is set to 0 outside of its domain of definition Ω) and ϕ_n represents the standard normal n -dimensional density function and $r(G) = n - I(g * \phi_n)$, where $I(g * \phi_n)$ is the Fisher information.

So far, the only known explicit solutions for the asymptotically minimax estimation of a bounded Gaussian vector were obtained on the sphere, an orthosymmetric domain, or on rectangles, which are products of orthosymmetric domains. However, much interesting statistical research has also occurred in parameter estimation under positivity or order constraints (for an overview of this subject, see Robertson, Wright, and Dykstra (1988)). For example, the estimation of a Gaussian vector under order constraints on the components has been shown to be a necessary part of the solution to some statistical problems in toxicology experiments (Schoenfeld (1986)).

The prior information about ω for such a problem could consist of it being contained in some domain $\Omega \subset \mathbb{R}^n$ such as the domain given by the inequalities $\omega_1 \geq \omega_2 \geq \dots \geq \omega_n$ or $\omega_1 \geq \omega_2 \geq \dots \geq \omega_n \geq 0$. Such inequalities arise in many situations in order restricted inference (Robertson *et al.* (1988)). These two domains are well known as the standard fundamental domains (chambers) for certain finite groups generated by reflections (the groups of permutations, respectively, signed permutations of the coordinates). So, with the same effort, we may ask the question about the fundamental chamber C of an arbitrary finite reflection group W . We cannot handle these domains directly, one reason being that Theorem 1 works smoothly only for bounded domains Ω . But what can be done, and is still useful, is to consider $\Omega = C \cap B$ where B is the unit ball of $\mathbb{R}^n$. Using a symmetry argument used before by Berard (1980) and Berard and Besson (1980) in the case of the intersection of the sphere with a Weyl chamber (see also Urakawa (1982, 1984), we show how to obtain from the class of eigenfunctions on the ball the principal eigenfunctions of Ω . Then the asymptotically least favourable distributions can be derived and second order asymptotically minimax estimates for the associated minimax problem on $t\Omega$ can also be constructed in a manner analogous to Bickel's work (1981) in the one dimensional case.

2 Asymptotically least favorable prior distributions on a bounded domain equal to the intersection of a ball and the fundamental chamber of a reflection group

Here our goal is to study some of the important domains for order restricted statistical inference and to obtain asymptotically least favorable prior distributions for such domains. The four chosen, each of which is a subset of the unit ball B in $\mathbb{R}^n$, are as follows:

$$\begin{aligned} (i) \quad \mathcal{S} &= \{x \in \mathbb{R}^n : |x| \leq 1, x_1 \geq 0\} \\ (ii) \quad \mathcal{Q} &= \{x \in \mathbb{R}^n : |x| \leq 1 \text{ and } x_1, \dots, x_n \geq 0\} \quad \text{and} \\ (iii) \quad \mathcal{A} &= \{x \in \mathbb{R}^n : |x| \leq 1 \text{ and } x_1 \geq x_2 \geq \dots \geq x_n\} \\ (iv) \quad \mathcal{B} &= \{x \in \mathbb{R}^n : |x| \leq 1 \text{ and } x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}, \end{aligned} \tag{4}$$

In order to achieve this, we first study the eigenfunctions on the unit ball; then a general theorem about the eigenfunctions of a domain equal to the intersection of a ball and the fundamental chamber of a reflection group is proved. The results for the domains given in (4) and for other interesting domains of this type then follow as special cases.

In the following, we use the notation $j_\nu(r) = r^{-\nu} J_\nu(r)$, where J_ν is the standard Bessel function with index ν . We denote by $k_{\nu,1} < k_{\nu,2} < \dots$ the positive zeroes of J_ν (hence also of j_ν). We recall that $j_\nu(r)$ is an entire analytic function of r , for every ν (Bateman II, p. 4).

In order to obtain asymptotically minimax estimates for $t\Omega$ where $\Omega = C \cap B$, we must have an expression for the principal eigenfunction h and the corresponding eigenvalue λ for Ω . The key idea of the proof is a symmetry argument also used in a similar context for

the sphere by Berard (1980) and Berard and Besson (1980) (see also Urakawa 1982, 1984). It can be summarized as follows.

Let W denote a finite group generated by reflections of $\mathbb{R}^n$. Then we have a set of hyperplanes $\mathcal{Z}$, invariant under W , such that the orthogonal reflections s_H with $H \in \mathcal{Z}$ generate W . The connected components of $\mathbb{R}^n \setminus \bigcup_{H \in \mathcal{Z}} H$, the complement of all the $H \in \mathcal{Z}$, are called fundamental domains or chambers for W .

Each of the bounded domains $\mathcal{S}$, $\mathcal{Q}$, $\mathcal{A}$, $\mathcal{B}$ arises in this way as the intersection of the ball B with a fundamental chamber C of a certain W . Now the eigenfunctions of the Dirichlet problem for the ball B can all be written as products of Bessel functions and harmonic polynomials. Then one can prove that the W -anti-invariant elements of $L^2(B)$ restricted to C give exactly $L^2(C \cap B)$. Using some known facts about zeroes of Bessel functions the problem now reduces to finding the anti-invariant harmonic polynomial of least degree.

Theorem 2 *Up to an arbitrary constant factor the solutions of the eigenvalue problem for the n dimensional ball with Dirichlet boundary condition are exactly:*

$$f(x) = j_\nu(|x|k_{\nu,m})h_l(x) \quad (5)$$

where $\nu = l + \frac{n-2}{2}$, $l = 0, 1, 2, \dots$, $m = 1, 2, \dots$, where $h_l \in \mathcal{H}_l$ is a homogeneous harmonic polynomial of degree l , and where $k_{\nu,m}$ is the m^{th} zero of J_ν . The corresponding eigenvalue is $k_{\nu,m}^2$.

The principal eigenfunction is given by $l = 0, m = 1$, that is:

$$j_{\frac{n-2}{2}}(|x|k_{\frac{n-2}{2},1}). \quad (6)$$

Proof: For the proof of (5) see Theorem 2.66, p. 107 of Folland (1995).

The principal eigenfunction is the one given by $l = 0, m = 1$ as in (6) because $k_{\nu,1} < k_{\nu+1,1} < k_{\nu,2} < k_{\nu+1,2} \dots$ (see Watson, p. 479 or Bateman II, p. 59).

□

Now let us consider the general case. Henceforth W will denote a finite group generated by reflections in $\mathbb{R}^n$, C a fundamental chamber for W . More precisely, let $\mathcal{Z}$ be a finite set of hyperplanes in $\mathbb{R}^n$, such that the orthogonal reflection s_H with respect to each $H \in \mathcal{Z}$ preserves the set $\mathcal{Z}$, and let W be the group generated by all the s_H ($H \in \mathcal{Z}$). W is a finite Coxeter group; canonically associated with W are natural numbers $m_1, \dots, m_n$ called the exponents of W (cf. Bourbaki, Ch. 5, § 6, No 2). The connected components of $\mathbb{R}^n \setminus \bigcup_{H \in \mathcal{Z}} H$ are called chambers. It is well known (Bourbaki, Ch. 5, § 3, Thm. 1 (p. 74)) that W is simply transitive on the set of chambers. Let us fix one of them and call it C .

We choose a unit vector e_H orthogonal to H for each H so that $(x|e_H) > 0, \forall x \in C$. We note that,

$$\Delta(f \circ w) = (\Delta f) \circ w \quad (7)$$

for all $w \in W$ and all functions f , where Δ denotes the Laplacian. This is clear since each w is orthogonal (since all s_H are orthogonal). In particular, if f is harmonic, so is $f \circ w$. We say that a function f is invariant if $f \circ w = f$ ($\forall w \in W$) and anti-invariant if $f \circ w = (\det w)f$ ($\forall w \in W$).

Let Π be defined, $\forall x \in \mathbb{R}^n$, by:

$$\Pi(x) = \prod_{H \in \mathcal{Z}} (x|e_H) \quad (8)$$

Theorem 3 *Let W , C and Π be defined as above and let $\Omega = B \cap C$, where B is the unit ball in $\mathbb{R}^n$. Then the least favorable prior density on Ω is given by*

$$g_\Omega = c_\Omega (j_\nu(|x|k_{\nu,1}))^2 (\Pi(x))^2 \quad x \in \Omega \text{ otherwise} \quad (9)$$

with $\nu = N + \frac{n-2}{2}$ (where $N = |\mathcal{Z}|$, the number of planes in $\mathcal{Z}$ and also equals the degree of Π) and $c_\Omega = 2^{2N} \pi^{-n/2} \Gamma(\nu + 1) (J_{\nu+1}(k_{\nu,1}))^{-2} \left(\prod_{i=1}^n (m_i!) \right)^{-1}$ where $m, \dots, m_n$ are the exponents of W .

Proof: The first part of the proof, the determination of the eigenfunctions of Ω can be found in Berard (1980) and Berard and Besson (1980) together with Urakawa (1982, 1984). For completeness we include it here.

We define operator *Alt* on functions by:

$$(Alt f)(x) = \frac{1}{|W|} \sum_{w \in W} (\det w) f(wx)$$

where $|W| = \#$ of elements in W . Considered on $L^2(B)$ it is the orthogonal projection onto the subspace $L^2(B)^{Alt}$ of anti-invariant elements; (this is clear since it is idempotent and self-adjoint and leaves fixed all anti-invariant f 's). *Alt* also preserves each $\mathcal{H}_l$ by (7). Therefore it also preserves each $L^2(B)_{l,m}$, the subspace of $L^2(B)$ spanned by the functions (5) for fixed l, m . We denote the images of *Alt* in these spaces by $L^2(B)_{l,m}^{Alt}$ and $\mathcal{H}_l^{Alt}$. We know that: $L^2(B) = \oplus_{l=0}^\infty \oplus_{m=1}^\infty L^2(B)_{l,m}$. It follows therefore that: $L^2(B)_{l,m}^{Alt} = \oplus_{l=0}^\infty \oplus_{m=1}^\infty L^2(B)_{l,m}^{Alt}$.

We claim that the restrictions to Ω of the functions in $L^2(B)_{l,m}^{Alt}$ ($l = 0, 1, \dots; m = 1, 2, \dots$) span $L^2(\Omega)$ (and are still orthogonal to each other in $L^2(\Omega)$ for l, m different). For this it is now enough to see that the restriction map (from B to Ω) is (up to a scalar) a Hilbert space isomorphism of $L^2(B)^{Alt}$ onto $L^2(\Omega)$. For this we note that (by the simple transitivity of W on the chambers)

$$\begin{aligned}
\int_B fg &= \sum_{w \in W} \int_{w\Omega} fg \\
&= \sum_{w \in W} \int_{\Omega} f(w^{-1}x)g(w^{-1}x) dx \\
&= |W| \int_{\Omega} f(x)g(x) dx,
\end{aligned}$$

and that the restriction map is surjective, since for any φ in $L^2(\Omega)$ we have the anti-invariant extension $\tilde{\varphi}$ on $L^2(B)$ defined uniquely by $\tilde{\varphi}(wx) = (\det w)\varphi(x)$ for $w \in W$ and $x \in \Omega$. We now have a complete description of the Dirichlet eigenfunctions of Ω ; they are given by (5) with the restriction that $h = h_l \in \mathcal{H}_l^{Alt}$.

We now only have to find the smallest l such that $\mathcal{H}_l^{Alt} \neq 0$ (and a function in this $\mathcal{H}_l^{Alt}$) and to determine the constant c_{Ω} .

By Bourbaki, Ch. 5, § 5, Prop. 5 (p. 113), Π , defined in (8), is the lowest degree anti-invariant (with respect to W) polynomial on $\mathbb{R}^n$. Here is a possibly simpler proof of this. To show Π is anti-invariant we have to show $\Pi \circ s_{H_0} = -\Pi$ ($\forall H_0 \in \mathcal{Z}$). Now

$$\Pi(s_{H_0}x) = \prod_{H \in \mathcal{Z}} (s_{H_0}x|e_H) = \prod_{H \in \mathcal{Z}} (x|s_{H_0}e_H).$$

As H runs through all hyperplanes in $\mathcal{Z}$, $s_{H_0}e_H$ will run through all e_H with possible sign changes. It follows that $\Pi \circ s_{H_0} = \pm\Pi$. To determine the sign it suffices to look at one special $x \in \mathbb{R}^n$. We take an x that is very close to the plane H_0 and far from all other planes in $\mathcal{Z}$ ($(x|e_{H_0})$ is small; $(x|e_H)$ is large for $H \neq H_0$). Then $s_{H_0}x$ is close to x , so $(s_{H_0}x|e_H)$ for $H \neq H_0$ has the same sign as $(x|e_H)$ but $(s_{H_0}x|e_{H_0}) = (x|s_{H_0}e_{H_0}) = -(x|e_{H_0})$. So $\Pi(s_{H_0}x) = -\Pi(x)$, and $\Pi \circ s_{H_0} = -\Pi$.

To show that Π is of lowest degree, following Bourbaki, we observe that any semi-invariant is 0 on every $H \in \mathcal{Z}$, hence divisible by $(x|e_H)$, hence divisible by Π .

In addition, it is well known that Π is harmonic. This can be seen as follows. Since Alt commutes with Δ and Π is anti-invariant, $\Delta\Pi = \Delta Alt(\Pi) = Alt(\Delta\Pi)$, so $\Delta\Pi$ is anti-invariant. But the degree of $\Delta\Pi$ is strictly less than $\deg(\Pi)$, hence $\Delta\Pi = 0$.

Now from (5) and from the inequalities for the $k_{\nu,m}$ we see that the principal eigenfunction is given by

$$|x|^{-\nu} J_{\nu}(|x|k_{\nu,1})\Pi(x)$$

Once the principal eigenfunction has been found it suffices to use Theorem 1 to determine the asymptotically least favorable prior density on Ω given by

$$g_{\Omega}(x) = c_{\Omega}|x|^{-2\nu} (J_{\nu}(|x|k_{\nu,1}))^2 (\Pi(x))^2.$$

In order to determine the constant c_Ω , since $\int g_\Omega(x)dx = 1$, we have, using polar coordinates, $x = rx'$, $x \in \Sigma$ with Σ denoting the unit sphere.

$$\begin{aligned} \frac{1}{c_\Omega} &= \int_{\Omega} (|x|^{-\nu} J_\nu(|x|k_{\nu,1}) \Pi(x))^2 dx \\ &= \frac{1}{|W|} \int_0^1 r (J_\nu(k_{\nu,1}r))^2 dr \int_{\Sigma} (\Pi(x'))^2 d\sigma(x'). \end{aligned}$$

By Watson (1962, Section 18.1), the first integral on the right hand side equals $\frac{1}{2} (J_{\nu+1}(k_{\nu,1}))^2$. We denote the second integral by K_Ω and we consider

$$L_\Omega = \int_{\mathbb{R}^n} e^{-|x|^2/2} (\Pi(x))^2 dx.$$

Using polar coordinates again, we obtain

$$L_\Omega = 2^\nu \Gamma(\nu + 1) K_\Omega.$$

On the other hand, Macdonald (1980) pointed out that Π being a harmonic polynomial, the integral L_Ω is equal to the norm computed by R. Steinberg and written down in the Appendix of Harder (1971) namely

$$L_\Omega = (2\pi)^{n/2} \frac{|W|}{2^N} \prod_{l=1}^n (m_l!)$$

where $m_1, \dots, m_n$ are the exponents of W .

Thus,

$$\frac{1}{c_\Omega} = \frac{\pi^{n/2}}{2^{2N} \Gamma(N + \frac{n}{2})} (J_{\nu+1}(k_{\nu,1}))^2 \prod_{l=1}^n (m_l!)$$

with $N + \frac{n}{2} = \nu + 1$.

□

Remark The exponents m_l are given in Bourbaki (1969) for W arising from a root system in the tables on pages 250–276 and for W of type H_3 or H_4 on pages 231–232. In the only remaining case, W of type $I_2(p)$, we have trivially $m_1 = 1, m_2 = p - 1$.

Let us now consider the special domains given in equation (4). For an n -vector $x = (x_i, \dots, x_n)$ and a positive t let us introduce the following notation:

$$x_t = (x_{1,t}, x_{2,t}, \dots, x_{n,t})$$

where $x_{i,t} = \frac{x_i}{t}$ ($1 \leq i \leq n$). Note that if $x_t \in \Omega$ then $x \in t\Omega$.

Corollary 1 *The asymptotically least favorable prior densities for $t\mathcal{Q}$, $t\mathcal{A}$ and $t\mathcal{B}$ where $\mathcal{Q}$, $\mathcal{A}$ and $\mathcal{B}$ are defined in (4) are given respectively by*

$$\begin{aligned} g_{t\mathcal{Q}}(x_1) &= c_{\mathcal{Q}} t^{-n} \left[|x_t|^{-\frac{3n-2}{2}} J_{\frac{3n-2}{2}} \left(k_{\frac{3n-2}{2},1} |x_t| \right) x_{1,t} \cdots x_{n,t} \right]^2 \\ g_{t\mathcal{A}}(x) &= 2^{-\frac{n(n-1)}{2}} c_{\mathcal{A}} t^{-n} \left[|x_t|^{-\frac{n^2-2}{2}} J_{\frac{n^2-2}{2}} \left(k_{\frac{n^2-2}{2},1} |x_t| \right) \prod_{i < j} (x_{i,t} - x_{j,t}) \right]^2 \\ g_{t\mathcal{B}}(x) &= 2^{-n(n-1)} c_{\mathcal{B}} t^{-n} \left[|x_t|^{-\frac{n-2}{2}} J_{\frac{2n^2+n-2}{2}} \left(k_{\frac{2n^2+n-2}{2},1} |x_t| \right) \prod_{i=1}^n x_{i,t} \prod_{i < j} (x_{i,t} - x_{j,t})^2 \right]^2. \end{aligned} \quad (10)$$

Proof: By Theorem 1 it suffices to use Theorem 3 and to determine the principal eigenfunction for each domain.

The case of $\mathcal{A}$, for example, is the special case where $\mathcal{Z}$ is the set of all hyperplanes $x_i - x_j = 0$, ($i > j$). Here $N = |\mathcal{Z}| = \frac{n(n-1)}{2}$, and $\Pi(x) = 2^{-N/2} \prod_{i < j} (x_i - x_j)$. (So $\nu = \frac{n^2-2}{2}$.)

The case where $\mathcal{B} = \{x \in \mathbb{R}^n; |x| \leq 1 \text{ and } x_1 \geq x_2 \geq \cdots \geq x_n > 0\}$ corresponds to $\mathcal{Z}$, the set of all $x_i - x_j = 0$, $x_i + x_j = 0$, ($i > j$), $x_i = 0$. So $N = |\mathcal{Z}| = n^2$, and $\Pi(x) = 2^{-\frac{n(n-1)}{2}} \prod_{i=1}^n x_i \prod_{i < j} (x_i^2 - x_j^2)$.

Analogous arguments work for $\mathcal{Q}$ and $\mathcal{S}$. The explicit values of the constants $c_{\mathcal{S}}$, $c_{\mathcal{Q}}$, $c_{\mathcal{A}}$ and $c_{\mathcal{B}}$ can be obtained in a similar fashion by specifying the parameters in the formula for c_{Ω} given in Theorem 3.

Remark *Chapter VI, § 4 of Bourbaki (1969) contains a list of all possible W 's (and C 's). For each case we can write down explicit formulas for $g_{t\Omega}(x)$ where $\Omega = C \cap B$.*

3 Construction of asymptotically minimax estimators

As in Section 2, we consider domains $\Omega = C \cap B$ where C is the fundamental chamber for a finite group generated by reflections and B is the unit ball in $\mathbb{R}^n$. We denote by λ the smallest eigenvalue of the Dirichlet eigenvalue problem (2) for Ω . Our goal is to construct a family of estimators $\delta_t^*(x)$ associated to the domains $t\Omega$ ($t > 0$) which is asymptotically minimax in the sense that

$$\sup_{\omega \in t\Omega} \left| R(\omega, \delta_t^*) - \left(n - \frac{4\lambda}{t^2} \right) \right| = o\left(\frac{1}{t^2}\right). \quad (11)$$

The essential idea of the proof that an estimator δ_t^* satisfies (11) involves expressing the estimator as

$$\delta_t^*(x) = x - \psi^{(t)}(x) \quad (12)$$

and substituting this expression in Stein's unbiased estimate of the risk:

$$R(\omega, \delta_t^*) = n - \int_{\mathbb{R}^n} (2 \operatorname{div} \psi^{(t)}(x) - |\psi^{(t)}(\omega)|^2) \phi_n(x - \omega) dx, \quad (13)$$

where div denotes divergence, $|\psi^{(t)}(\omega)|$ denotes the norm in $\mathbb{R}^n$ and $\phi_n(x)$, the standard normal n -dimensional density function.

Clearly the Bayes estimator

$$\delta_t(x) = x + \operatorname{grad} \log(g_t * \phi_n)$$

associated to the “approximately least favorable density δ_t defined in Theorem 1 and explicitly computed in Theorem 3 does not have the property defined by (11), because as noted by Bickel (1981) and Berkin and Levit (1981), since u_Ω vanishes on the boundary of Ω , the behaviour of $\nabla u_\Omega / u_\Omega$ is unstable near the boundary. This renders it necessary to construct δ_t^* , a suitable modification of δ_t , by considering an appropriate neighbourhood $\Omega^{(t)}$ of Ω and analogously to Bickel's work (1981) in the one dimensional case, modifying $v_\Omega(x)$ from Theorem 1 inside and outside $\Omega^{(t)}$ in such a way that we can prove that the following error terms are uniformly small:

$$R_{1t} = \max_{\omega \in t\Omega} \left| \int_{t\Omega^{(t)}} \left(2 \operatorname{div} \psi^{(t)}(x) - \left\| \psi^{(t)}(x) \right\|^2 - \frac{4\lambda}{t^2} \right) \phi_n(x - \omega) dx \right| = o\left(\frac{1}{t^2}\right), \quad (14)$$

$$R_{2t} = \max_{\omega \in t\Omega^c} \left| \int_{(t\Omega^{(t)})^c} \left(2 \operatorname{div} \psi^{(t)}(x) - \left\| \psi^{(t)}(x) \right\|^2 - \frac{4\lambda}{t^2} \right) \phi_n(x - \omega) dx \right| = o\left(\frac{1}{t^2}\right). \quad (15)$$

We proceed to construct $\Omega^{(t)}$ and δ_t^* as follows. In order to simplify notation, let $e_i (1 \leq i \leq N)$ denote the vectors e_H of (8) arranged in some order. Let us choose a unit vector $u \in C$ so that $(u|e_i) \geq 0 \forall i$ and let us choose k such that $k \geq \frac{1}{\min_i (u|e_i)}$. Let $a_t < 1$ satisfying $a_t \downarrow 0$ and $ta_t \rightarrow \infty$. Now let b_t be such that $0 < b_t < \frac{a_t}{2(1 + (k+1)a_t)}$

and $b_t^{-4} \int_{|x| > \frac{ta_t}{2}} \phi_n(x) dx \rightarrow 0$.

Let $|x| = r$ and let

$$\begin{aligned} \Omega &= \{x \in C : |x| \leq 1\} \\ \Omega_t &= \{x \in C : r < 1 - b_t \text{ and } (x|e_i) > b_t \text{ } (1 \leq i \leq N)\} \\ \Omega^{(t)} &= (1 + (k+1)a_t)\Omega_t - a_t k u. \end{aligned}$$

It is easy to see that, if $\omega \in \Omega$ and $x \in (\Omega^{(t)})^c$, then $|x - \omega| > \frac{a_t}{2}$.

We define $\bar{\psi}$, for $x \in \Omega$ only, by

$$\bar{\psi}(x) = \frac{-\nabla g_\Omega(x)}{g_\Omega(x)},$$

where g_Ω is as in Theorem 3. From $\Delta g_\Omega(x) = -\lambda g_\Omega(x)$ it follows that

$$2 \operatorname{div} \bar{\psi}(x) - |\bar{\psi}(x)|^2 = 4\lambda \quad (16)$$

for all x in Ω .

Using the auxiliary functions

$$\begin{aligned} \varphi_0(y) &= k_{\nu,1} \frac{j'_\nu(k_{\nu,1}y)}{j_\nu(k_{\nu,1}y)} & (0 \leq y < 1) \\ \varphi(y) &= \frac{1}{y} & (0 < y) \end{aligned}$$

the explicit formula (9) gives

$$\bar{\psi}(x) = -2\varphi_0(r) \frac{x}{r} - 2 \sum_i \varphi((x|e_i)) e_i.$$

We modify $\bar{\psi}(x)$ near the boundary of Ω and extend its definition to $\mathbb{R}^n$ by first setting

$$\begin{aligned} \varphi_{0,t}(y) &= \begin{cases} \varphi_0(y) & \text{if } 0 \leq y \leq 1 - b_t \\ \varphi_0(1 - b_t) + \varphi'_0(1 - b_t)(y - (1 - b_t)) & \text{if } y > 1 - b_t \end{cases} \\ \varphi_t(y) &= \begin{cases} \varphi(y)(= \frac{1}{y}) & \text{if } y \geq b_t \\ \varphi(b_t) + \varphi'(b_t)(y - b_t) & \text{if } y < b_t \end{cases} \end{aligned}$$

and then for $x \in \mathbb{R}^n$ and $r = |x|$

$$\begin{aligned} \psi_{0,t}(x) &= -2\varphi_{0,t}(r) \frac{x}{r}, \\ \psi_{i,t}(x) &= -2\varphi_t((x|e_i)) e_i, \\ \psi_t(x) &= \psi_{0,t}(x) + \sum_i \psi_{i,t}(x). \end{aligned}$$

Finally, we define,

$$\psi^{(t)}(x) = \frac{1}{(1 + (k+1)a_t)t} \psi_t \left(\frac{1}{(1 + (k+1)a_t)t} (x + a_t k u) \right). \quad (17)$$

Theorem 4 *An asymptotically minimax estimator $\delta_t^*(x)$ which satisfies (11) can be defined by letting*

$$\delta_t^*(x) = x - \psi^{(t)}(x) \quad (18)$$

where $\psi^{(t)}$ is defined in (17).

Proof:

We will use the following properties in the proof.

$$(a) \quad 2 \operatorname{div} \psi^{(t)}(x) - \|\psi^{(t)}(x)\|^2 = \frac{4\lambda}{(1 + (k+1)a_t)^2 t^2} \text{ for } x \in t\Omega^{(t)}. \text{ This follows from (12)}$$

and from the fact that $\psi_t = \bar{\psi}$ on Ω_t .

$$(b) \quad 2 \operatorname{div} \psi_t(x) - \|\psi_t(x)\|^2 \leq \frac{c}{b_t^4} (r^2 + 1) \text{ for some } c \text{ independent of } x \text{ and } t.$$

To see this, we first get by direct computation

$$\begin{aligned} \operatorname{div} \psi_{ot}(x) &= -2 \left(\varphi'_{ot}(r) + \varphi_{ot}(r) \frac{n-1}{r} \right), \\ \operatorname{div} \psi_{it}(x) &= -2\varphi'_t((x|e_i)), \\ |\psi_{ot}(x)|^2 &= 4\varphi_{ot}(r)^2, \\ |\psi_{it}(x)|^2 &= 4\varphi_t((x|e_i))^2. \end{aligned}$$

Then we observe that the functions $\varphi'_t(y), \varphi_t(y)$ (for $y \in \mathbb{R}$) and $\varphi'_{ot}(y), \varphi_{ot}(y)$ (for $y \geq 0$) are all majorized by

$$\frac{c}{b_t^2} (y + 1)$$

for some constant c . This is obvious for φ'_t and φ_t ; to prove it for φ'_{ot} we note that, $k_{\nu,1}$ being a simple zero and the smallest positive zero of j_ν , we have, for $0 < y < 1$, $\varphi'_0(y) = -\frac{1}{(1-y)^2} + a(y)$ for a smooth function $a(y)$ on $[0, 1]$. From this the estimate for φ'_{ot} and then for φ_{ot} follows and (b) is then an immediate consequence.

Now, to prove the theorem, we have to show that assertions (14) and (15) are true. It is easy to see that

$$\left(\frac{4\lambda}{(1 + (k+1)a_t)^2 t^2} - \frac{4\lambda}{t^2} \right) \int_{t\Omega^{(t)}} \phi_n(x - \omega) dx = o\left(\frac{1}{t^2}\right),$$

and this implies (14) by (a). So it is enough to prove R_{2t} in (15) is uniformly small. In other words, we have to prove that

$$\begin{aligned} \max_{\omega \in t\Omega} \left| \int_{(t\Omega^{(t)})^c} \left\{ \frac{1}{(1 + (k+1)a_t)^2} \left(2 \operatorname{div} \psi_t - \|\psi_t\|^2 \right) \right. \right. \\ \left. \left. \left(\frac{1}{(1 + (k+1)a_t)t} (x + a_t k u) \right) - 4\lambda \right\} \phi_n(x - \omega) dx \right| = o(1) \end{aligned} \quad (19)$$

As previously observed if $x \in (t\Omega^{(t)})^c$ and $\omega \in t\Omega$, then $|x - \omega| > \frac{ta_t}{2}$. Changing the variable $z = x - \omega$, and applying (b), we find that the left hand side of (19) is majorized by

$$\frac{c''}{b_t^4} \int_{|z| > \frac{ta_t}{2}} (|z|^2 + c) \phi_n(z) dz.$$

and this is really $o(1)$ by the hypotheses about a_t and b_t . □

Acknowledgements

The first author was partially supported by NSF of the USA and by a PSC-CUNY award. The second author would like to acknowledge the partial support of NSERC of Canada and FCAR of Québec for this research. Both authors are grateful to the referees for suggestions and indicating certain references.

References

- BERARD, P., (1980), Spectres et groupes cristallographiques I: Domaines euclidiens, *Inventiones Mathematicae*, **58**, 179–199.
- BERARD, P., and BESSON, G., (1980), Spectres et groupes cristallographiques II: Domaines spheriques, *Annales de l'Institut Fourier*, **30**, 237–248.
- BERGER, J.O. (1985), *Statistical decision theory and Bayesian analysis*, Second Edition, Springer-Verlag, New York.
- BERKIN, P.E. and LEVIT B. Ya. (1981), Second-order asymptotically minimax estimates for the mean of a normal population, *Problemy Peredachi Informatsii*, **163**, (1980) 60–79. Translation: *Problems of Information Transmission*, 212–229.
- BICKEL, P.J. (1981), Minimax estimation of the mean of a normal distribution when the parameter space is restricted, *Ann. Statist.*, **9**, 1301–1309.
- BOURBAKI, N. (1969), *Éléments de mathématique*, Fascicule XXXIV, Groupes et algèbres de Lie, Chapitres IV, V et VI, Hermann, Paris.
- BROWN, L.D. (1979), *Statistical Decision Theory, Mimeographed Notes*, Cornell University, Ithaca, N.Y.
- BROWN, L.D. and GAJEK, L. (1990), Information inequalities for the Bayes risk, *Ann. Statist.*, **18** 1578–1594.
- CASELLA, G. and STRAWDERMAN, W. (1981), Estimating a bounded normal mean, *Ann. Statist.*, **9** 868–876.
- DONOHU, D.L., LIU, R.C. and MACGIBBON, B. (1990), Minimax risk over hyperrectangles and implications, *Ann. Statist.*, **18** 1416–1437.
- ERDÉLYI, A. (1953–1955), The Bateman Manuscript Project (A. Erdélyi, ed.) *Higher Transcendental Functions* (3 volumes), McGraw-Hill, New York.
- FERGUSON, T.S. (1967), *Mathematical Statistics, A Decision Theoretic Approach*, Academic, New York.
- FOLLAND, G.B. (1995), *Introduction to Partial Differential Equations*, 2nd edition, Princeton University Press.
- GILBARG, D. and TRUDINGER, N.S. (1983), *Elliptic Partial Differential Equations of Second Order*, 2nd ed. Springer-Verlag, Berlin-Heidelberg-New York.

- HARDER, G. (1971), A Gauss-Bonnet formula for discrete arithmetically defined groups, *Ann. Sci. Ec. Norm. Super.*, **4** 409–445.
- JOHNSTONE, I.M. and MACGIBBON, K.B. (1992), Minimax estimation of a constrained Poisson vector, *Ann. Statist.*, **20** 807–831.
- JOHNSTONE, I.M. and MACGIBBON, K.B. (1993), Asymptotically minimax estimation of a constrained Poisson vector via polydisc transforms, *Annales de l'Institut Henri Poincaré - Probabilité et Statistiques*, **29** 289–319.
- KEMPTHORNE, P.J. (1987), Numerical specification of discrete least favorable prior distributions, *SIAM J. Sci. Statist. Computing* **8** 171–184.
- LEVIT, B. Ya. (1980), On asymptotic minimax estimates of the second order, Translation: *Theor. Prob. Appl.*, **25** 552–568.
- LEVIT, B. Ya. (1982), Minimax estimation and positive solutions of elliptic equations, Translation: *Theor. Prob. Appl.*, **27** 563–586.
- LEVIT, B. Ya. 1985a, Second order asymptotic optimality and positive solutions of Schrödinger's equation, Translation: *Theor. Prob. Appl.*, **30** 333–363.
- LEVIT, B. Ya. (1985b), Evaluation of the minimax risk, *Fourth International Vilnius Conf. Prob. Theory Math. Stat.*, **4** Vilnius, 181–183.
- LEVIT, B. Ya. (1986), On second order admissibility in simultaneous estimation, *Proc. 1st. World Congress Bernoulli Soc.. Tashkent*, U.S.S.R., Yu. A. PROHOROV and V.V. SAZONOV Eds., V.N.U. Sc. Press, Utrecht.
- MACDONALD, P.G. (1980), The volume of a compact Lie Group, *Inventiones Mathematicae*, **56** 93–95.
- ROBERTSON, T., WRIGHT, F.T. and DYKSTRA, R.L. (1988), *Order Restricted Statistical Inference*, John Wiley and Sons, New York.
- SCHOENFELD, D.A. (1986), Confidence bounds for normal means under order restrictions, with applications to dose-response curves, toxicology experiments, and low-dose extrapolation, *J. Amer. Statist. Assoc.*, **81** 186–195.
- URAKAWA, H. (1982), Bounded domains which are isospectral but not congruent, *Ann. Sci. Ec. Norm. Super.*, **IV(15)** 441–456.
- URAKAWA, H. (1984), Reflection groups and the eigenvalue problems of vibrating membranes with mixed boundary conditions, *Tohoku Math. J.*, **II(36)**, 175–183.
- WALD, A. (1950), *Statistical Decision Functions*, John Wiley, New York.
- WATSON, G. (1962), *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge.
- ZINZIUS, E. (1981), Minimaxschätzer für den Mittelwert v einer normalverteilten Zufallsgröße mit bekannter Varianz bei vorgegebener oberer und unterer Schranke für v . *Math. Operationsforsch. Statist. Ser. Statist.*, **12** 551–557.