

UNIVERSITAT DE BARCELONA

GLOBAL EFFICIENCY

by

J.M. Corcuera and J.M. Oller

AMS 1980 Subjects Classifications: 62F10, 62B10, 62A99



Mathematics Preprint Series No. 211

July 1996

Global efficiency

J. M. Corcuera and J. M. Oller
Department of Statistics
University of Barcelona
08028 Barcelona, Spain

June 18, 1996

Abstract

In this paper the global behaviour of an estimator is studied in framework of Intrinsic Analysis, [7]. Two indices of performance of an estimator in a bounded region are analyzed: the average of the intrinsic risk (the loss function is the squared Rao distance) and the maximum risk.

The Riemannian volume, provided by the Fisher metric on the manifold associated with the parametric model, allows us to take an average of the intrinsic risk. Cramér–Rao type integral inequalities for the integrated mean squared Rao distance of estimators are derived using variational methods, extending the work of Čencov, [3]. Additionally, lower bounds for the maximum risk are also derived, by using integral expressions.

AMS 1980 subjects classifications: 62F10, 62B10, 62A99.

Keywords and phrases: Riemannian metric, information metric, Rao distance, intrinsic analysis, Cramér–Rao type integral lower bounds, minimax risk.



1 Introduction

The purpose of this paper is to give lower bounds of two indices of the global behaviour of an estimator in a region of the parameter space: the average of the risk and the maximum risk. In this introduction we describe the framework where we are going to work and some results about local bounds of the risk that will be useful later to derive global bounds. Previous work in this area can be found in Prakasa Rao [8] and Čencov [3].

1.1 The framework

Let $\{(\chi, \mathbf{a}, P_\theta) ; \theta \in \Theta\}$ be a parametric statistical model, where Θ , the parameter space, is an n -dimensional C^∞ real manifold. Usually Θ is an open set of \mathbb{R}^n and in this case it is customary to use the same symbol, θ , to denote points and coordinates.

We suppose the map $\theta \mapsto P_\theta$ to be one-to-one and we consider the set of all probability measures in the statistical model, M , with the n -dimensional C^∞ real manifold structure induced by this map. Let us denote this manifold by (M, \mathfrak{A}) , where \mathfrak{A} is the atlas induced by the parametrizations, that is, the coordinates in the parameter space.

In the dominated case, which we shall assume hereafter, the probability measures can be represented by density functions. Then let us assume, for a fixed σ -finite reference measure μ , that $P_\theta \ll \mu$, $\forall \theta \in \Theta$ and denote by $p(\cdot; \theta)$ a density function with respect to μ , i.e., a certain version of the Radon-Nikodym derivative $dP_\theta/d\mu$. Now, through the identification $P_\theta \mapsto p(\cdot; \theta)$, the points in M can be considered either densities or probability measures. Additionally, we assume certain regularity conditions:

1. (M, \mathfrak{A}) is a connected Hausdorff manifold.
2. When x is fixed, the real function on M , $\xi \mapsto p(x; \xi)$, is a C^∞ function.
3. For every local chart (W, θ) , the functions in x , $\partial \log p(x; \theta) / \partial \theta^i$ $i = 1, \dots, n$, are linearly independent, and belong to $L^\alpha(p(\cdot; \theta) d\mu)$ for a suitable $\alpha > 0$.
4. The partial derivatives of the required orders

$$\partial / \partial \theta^i, \quad \partial^2 / \partial \theta^i \partial \theta^j, \quad \partial^3 / \partial \theta^i \partial \theta^j \partial \theta^k, \quad \dots, i, j, k = 1, \dots, n,$$

and the integration with respect to $d\mu$ of $p(x; \theta)$ can always be interchanged.

When all these conditions are satisfied, for a version of the density function, we shall say that the parametric statistical model is *regular*.

In this context, given a sample size k , an *estimator* \mathcal{U} for the true density function (or probability measure) $p = p(\cdot; \theta) \in M$ of the statistical model is a measurable map

$$\mathcal{U} : \chi^k \mapsto M,$$

assuming that the probability measure on χ^k is $(P)_k(dx) = p_{(k)}(x; \theta) \mu_k(dx) = \prod_{i=1}^k p(x_i; \theta) \mu(dx_i)$.

1.2 Local bounds

Let $h_{\alpha\beta}$ be a Riemannian metric on M and $g_{\alpha\beta}$ the Fisher metric. Then consider the Levi-Civita connection associated with $h_{\alpha\beta}$ and

$$A = \exp_p^{-1}(\mathcal{U}), \quad B = E_p(\exp_p^{-1}(\mathcal{U})),$$

and estimators \mathcal{U} such that B is a C^∞ field on M . Let $\mathfrak{S}_p = \{\xi \in M_p, |\xi| = 1\}$, M_p being the tangent space at p ; for each $\xi \in \mathfrak{S}_p$ we define

$$\mathcal{C}_p(\xi) = \sup\{s > 0 : d(p, \gamma_\xi(s)) = s\},$$

where d is the Riemannian distance and γ_ξ is a geodesic defined in an open interval containing zero, such that $\gamma_\xi(0) = p$ and with tangent vector at p equal to ξ . Now, if we set

$$\mathfrak{D}_p = \{s\xi \in M_p : 0 \leq s < \mathcal{C}_p(\xi) ; \xi \in \mathfrak{S}_p\}$$

and

$$D_p = \exp_p(\mathfrak{D}_p),$$

it is known that \exp_p maps \mathfrak{D}_p diffeomorphically onto D_p (see Hicks [5]). We have

Theorem 1.1 (Riemannian Cramér-Rao lower bound) . *Let \mathcal{U} be an estimator for a sample size k , corresponding to an n -dimensional regular parametric family of density functions. Assume that the manifold M is simply connected and that $(P)_k \circ \mathcal{U}^{-1}(M \setminus D_p) = 0$, $\forall p \in M$. Let us assume that the mean*

squared Riemannian distance given by $h_{\alpha\beta}$, between the true density and an estimate, $E(d^2(\mathcal{U}, p))$, exists and that the covariant derivative of B can be obtained by differentiating under the integral sign. Then,

$$E(d^2(\mathcal{U}, p)) \geq \frac{\{\text{div}(B) - E(\text{div}(A))\}^2}{kc} + \|B\|^2,$$

where $c = \sum_{\alpha, \beta} h^{\alpha\beta} g_{\alpha\beta}$ and $\text{div}(\cdot)$ stands for the divergence operator.

Proof: Let C be any vector field. Then, applying the Cauchy–Schwartz inequality twice,

$$E(|\langle A - B, C \rangle|) \leq E(\|A - B\| \|C\|) \leq \sqrt{E(\|A - B\|^2)} \sqrt{E(\|C\|^2)},$$

where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote, respectively, the inner product and the norm defined on each tangent space.

Let $C(x; \theta) = \text{grad}(\log p_{(k)}(x; \theta))$, where $\text{grad}(\cdot)$ is the gradient operator. Taking expectations and using the repeated index convention,

$$E(\|C\|^2) = E(h_{\alpha\beta} h^{\beta\gamma} \partial_\gamma \log p_{(k)} h^{\alpha\lambda} \partial_\lambda \log p_{(k)}) = k h_{\alpha\beta} h^{\beta\gamma} h^{\alpha\lambda} g_{\gamma\lambda} = k h^{\gamma\lambda} g_{\gamma\lambda}$$

where $\partial_\alpha = \partial/\partial\theta^\alpha$. Furthermore, we also have

$$|E(\langle A, C \rangle)| = |E(\langle A - B, C \rangle)| \leq E(|\langle A - B, C \rangle|)$$

and

$$E(\|A - B\|^2) = E(\|A\|^2) - \|B\|^2.$$

Thus,

$$|E(\langle A, C \rangle)| \leq \sqrt{E(\|A\|^2) - \|B\|^2} \sqrt{kc},$$

but $\|A\|^2 = d^2(\mathcal{U}, p)$. Moreover,

$$\text{div}(B) = E(\text{div}(A)) + E(\langle A, C \rangle),$$

and the theorem follows. ■

Remarks. We can choose a geodesic spherical coordinate system with origin $\mathcal{U}(x)$; under this coordinate system, we have

$$\frac{\partial A^\alpha}{\partial \theta^\alpha} = -1 \quad \text{and} \quad \Gamma_{\alpha j}^\alpha A^j = -\rho \Gamma_{\alpha 1}^\alpha = -\frac{\partial \log \sqrt{g}}{\partial \rho} \rho,$$

where g is the determinant of the metric tensor. Then

$$\text{div}(A) = -1 - \rho \frac{\partial \log \sqrt{g}}{\partial \rho}.$$

Now we can use Bishop's comparison theorems (see Chavel [4, pp. 71-73]) to estimate $\frac{\partial \log \sqrt{g}}{\partial \rho}$.

In the Euclidean case,

$$\frac{\partial \log \sqrt{g}}{\partial \rho} = \frac{n-1}{\rho},$$

and thus $\text{div}(A) = -n$.

When the sectional curvatures are non positive, we obtain

$$\frac{\partial \log \sqrt{g}}{\partial \rho} \geq \frac{n-1}{\rho},$$

and therefore $\text{div}(A) \leq -n$.

Finally, when the supreme of the sectional curvatures, \mathcal{K} , is positive and the diameter of the manifold satisfies $d(M) < \pi/2\sqrt{\mathcal{K}}$, we have

$$\frac{\partial \log \sqrt{g}}{\partial \rho} \geq 0,$$

and then we obtain $\text{div}(A) \leq -1$.

In any case, $\text{div}(A) \leq -a$, with $a = n$ or $a = 1$, depending on the sectional curvature sign.

Corollary 1.2 . Suppose there is a global chart such that $h_{\alpha\beta} = \delta_{\alpha\beta}$; identifying the points with their coordinates, we have,

$$\text{M.S.E.}(\mathcal{U}) \geq \frac{(\text{div}(E(\mathcal{U})))^2}{kg_{\beta\beta}} + (\text{Bias}(\mathcal{U}))^2,$$

where *M.S.E.* is the mean squared error.

Proof: It follows straightforward by the previous theorem and the facts that d is the Euclidean distance, $A = \mathcal{U} - p$ and $\text{div}(A) = -n$. ■

Corollary 1.3 (Intrinsic Cramér-Rao lower bound) . *If $h_{\alpha\beta} = g_{\alpha\beta}$, we have*

$$E(\rho^2(\mathcal{U}, p)) \geq \frac{\{\text{div}(B) - E(\text{div}(A))\}^2}{kn} + \|B\|^2,$$

where ρ is the Rao distance.

Proof: If the Riemannian metric is the Fisher metric the distance is known by Rao distance and $c = g^{\alpha\beta}g_{\alpha\beta} = \delta_\alpha^\alpha = n$. ■

1.3 Global bounds

Whatever loss function is considered, it is well known that, in general, there is no estimator whose risk function is uniformly smaller than any other. Therefore, given an estimator, it seems reasonable, in order to measure its performance over a certain region of the statistical model, to compute the integral of the risk and then to divide this quantity by the Riemannian volume of the region considered. *In the following, we take the Rao distance as loss function and the Riemannian metric as the Fisher metric. This is the Intrinsic Analysis framework.*

Let $\mathcal{B} \subset M$ be a measurable subset, with $0 < V(\mathcal{B}) < \infty$, where V is the Riemannian measure. We denote the *Riemannian average of the mean squared Rao distance* by

$$\mathcal{R}_{\mathcal{U}}^2(\mathcal{B}) = \frac{\int_{\mathcal{B}} E(\rho^2(\mathcal{U}, p)) dV}{\int_{\mathcal{B}} dV};$$

the performance index obtained is a weighted average of the mean squared distance. This approach is compatible with a Bayesian point of view: a uniform prior with respect to the Riemannian volume is a kind of noninformative prior (see Jeffreys [6]). It can be shown (see Berger [2]) that, when the parameter space

is a locally compact topological group, this Riemannian volume is a left invariant Haar measure and it is unique up to a multiplicative constant. In any case, this volume is invariant under any group that leaves the parametric family of densities invariant. In the first part of the paper we derive lower bounds for this global index on balls of radius R .

Another way to measure the global behaviour of an estimator is to consider the maximum risk in a region of the parameter space. This is a minimax approach. The last part of the paper is devoted to obtaining lower bounds for the maximum risk.

2 Variational methods to obtain global bounds

As we shall show, variational methods can be used to obtain global bounds. A previous study in this direction can be found in Čencov [3]. The idea is to consider the integral of the local bounds for the Rao distance given above when the Riemannian metric is the Fisher metric and on a submanifold $W \subset M$ with boundary $\partial W \subset M$ that is

$$\mathcal{Y}(B) = \int_W \left\{ \|B\|^2 + \frac{1}{kn}(\operatorname{div}(B) + a)^2 \right\} dV,$$

where we take $a = n$ if the sectional curvatures are non-positive and $a = 1$ in the other case. The above functional depends only on B and we can attempt to find the C^∞ vector field B that minimizes it. Since the minimum we obtain is in a class of vector fields larger than that of C^∞ bias vector fields, this method gives a lower bound for the *average of the mean squared Rao distance*.

Lemma 2.1 . *The C^∞ field B minimizes the functional*

$$\mathcal{Y}(B) = \int_W \left\{ \|B\|^2 + \frac{1}{kn}(\operatorname{div}(B) + a)^2 \right\} dV$$

iff it verifies

$$B - \frac{1}{kn} \operatorname{grad}(\operatorname{div}(B)) = 0, \quad \forall p \in W,$$

$$\operatorname{div}(B) + a = 0, \quad \forall p \in \partial W,$$

(1)

and the minimum value is given by

$$\mathcal{Y}^* = \frac{a^2}{kn} \text{vol}(W) - \frac{a}{kn} \int_{\partial W} \|B^*\| d\sigma = \frac{a^2}{kn} \text{vol}(W) + \frac{a}{kn} \int_W \text{div}(B^*) dV, \quad (2)$$

where B^* verifies (1).

Proof: Consider the first variation $\delta\mathcal{Y}(B, \eta)$, where η is an arbitrary field. Then it is easy to see that

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{Y}(B + \epsilon\eta) - \mathcal{Y}(B)}{\epsilon} \equiv \delta\mathcal{Y}(B, \eta) = \int_W \left(2\langle B, \eta \rangle + \frac{2}{kn} \text{div}(\eta) (\text{div}(B) + a) \right) dV$$

and

$$\mathcal{Y}(B + \eta) - \mathcal{Y}(B) = \delta\mathcal{Y}(B; \eta) + \int_W \left\{ \|\eta\|^2 + \frac{1}{kn} (\text{div}(\eta))^2 \right\} dV,$$

thus the functional is strictly convex and the stationary point is a global minimum. Now, the condition $\delta\mathcal{Y}(B; \eta) = 0$ is equivalent to

$$\int_W \left\{ \langle B, \eta \rangle + \frac{1}{kn} \text{div}(\eta) (\text{div}(B) + a) \right\} dV = 0.$$

If we take into account that

$$\text{div}(fX) = f \text{div}(X) + \langle X, \text{grad}(f) \rangle, \quad (3)$$

we obtain

$$\begin{aligned} \frac{1}{kn} \text{div}(\eta) (\text{div}(B) + a) &= \text{div} \left(\frac{1}{kn} \eta (\text{div}(B) + a) \right) - \langle \text{grad} \left(\frac{1}{kn} (\text{div}(B) + a) \right), \eta \rangle \\ &= \frac{1}{kn} \text{div}(\eta (\text{div}(B) + a)) - \frac{1}{kn} \langle \text{grad}(\text{div}(B)), \eta \rangle. \end{aligned}$$

We are able to write the stationarity condition as

$$\int_W \left\{ \langle B, \eta \rangle + \frac{1}{kn} \text{div}(\eta (\text{div}(B) + a)) - \frac{1}{kn} \langle \text{grad}(\text{div}(B)), \eta \rangle \right\} dV = 0$$

and by the Gauss divergence theorem,

$$\int_W \left\langle B - \frac{1}{kn} \text{grad}(\text{div}(B)), \eta \right\rangle dV + \frac{1}{kn} \int_{\partial W} \langle (\text{div}(B) + a)\eta, \nu \rangle d\sigma = 0,$$

where $d\sigma$ is the Riemannian measure on ∂W . Then (1) follows from the fact that the previous equality must be verified for any η .

Now we see the second part of the proposition. By the first stationarity condition (1) and by (3), we obtain

$$\|B^*\|^2 = \frac{1}{kn} (\operatorname{div}(B^* \operatorname{div}(B^*)) - (\operatorname{div}(B^*))^2),$$

and, putting it in $\mathcal{Y}(B)$,

$$\begin{aligned} \mathcal{Y}^* &= \int_W \frac{1}{kn} (\operatorname{div}(B^* \operatorname{div}(B^*)) + a^2 + 2a \operatorname{div}(B^*)) dV \\ &= \frac{a^2}{kn} \operatorname{vol}(W) + \frac{1}{kn} \int_{\partial W} \langle B^* \operatorname{div}(B^*) + 2aB^*, \nu \rangle d\sigma. \end{aligned}$$

Now, by the second stationarity condition in (1),

$$\mathcal{Y}^* = \frac{a^2}{kn} \operatorname{vol}(W) + \frac{a}{kn} \int_{\partial W} \langle B^*, \nu \rangle d\sigma. \quad (4)$$

It is clear that $0 \geq \operatorname{div}(B^*) \geq -a$, then since $\operatorname{div}(B^*) = -a$ on ∂W and $B^* = \operatorname{grad}(\operatorname{div}(B^*))$ it turns out that $\langle B^*, \nu \rangle = -\|B^*\|$. Finally by the Gauss divergence theorem we obtain the second equality in (2). ■

Remarks. Note that the minimum value of $\mathcal{Y}(B)$ depends only on $\operatorname{div}(B^*)$, and that $f^* \equiv \operatorname{div}(B^*)$ verifies the partial differential equation

$$\Delta f = knf, \quad \text{with } f(p) = -a, \quad \forall p \in \partial W, \quad (5)$$

as is easy to check from (1).

We have solved this boundary-value problem in the case where $W = S_R$, a ball of radius R , and constant sectional curvatures \mathcal{K} .

Theorem 2.2 . *When the parametric statistical model is a manifold of constant sectional curvature \mathcal{K} , we have the following lower bound for the average of the mean squared Rao distance, in balls of radius R such that $|\mathcal{K}|S_{\mathcal{K}}^2(R) < 1$:*

$$\mathcal{R}_U^2(S_R) \geq \frac{a^2}{kn} \left(1 - \frac{f'(R)S_{\mathcal{K}}^{n-1}(R)}{knf(R) \int_0^R S_{\mathcal{K}}^{n-1}(r) dr} \right), \quad (6)$$



where

$$f(R) = a_0 \sum_{j=0}^{\infty} \frac{\prod_{s=1}^j \{kn + 2\mathcal{K}(s-1)(n+2s-3)\}}{j! \left(\frac{n}{2}\right)_j 4^j} \mathbf{S}_{\mathcal{K}}^{2j}(R)$$

and

$$\mathbf{S}_{\mathcal{K}}(t) = \begin{cases} \frac{\sin(\sqrt{\mathcal{K}}t)}{\sqrt{\mathcal{K}}} & \text{if } \mathcal{K} > 0, \\ t & \text{if } \mathcal{K} = 0, \\ \frac{\sinh(\sqrt{-\mathcal{K}}t)}{\sqrt{-\mathcal{K}}} & \text{if } \mathcal{K} < 0. \end{cases} \quad (7)$$

Proof: By symmetry and uniqueness, the solution of the boundary problem in S_R ,

$$\Delta f = knf, \quad \text{with } f(p) = -a \quad \forall p \in \partial S_R, \quad (8)$$

depends only on the distance to the center of S_R . Then, taking geodesic spherical coordinates (r, u) with origin in the center of S_R , since $\sqrt{g}(r, u) = \mathbf{S}_{\mathcal{K}}^{n-1} \Omega(u)$ (see Appendix), we have

$$\Delta f = \frac{1}{\sqrt{g}} \frac{d}{dr} \left(\sqrt{g} \frac{d}{dr} f \right) = \frac{1}{\mathbf{S}_{\mathcal{K}}^{n-1}} \frac{d}{dr} \left(\mathbf{S}_{\mathcal{K}}^{n-1} \frac{d}{dr} f \right).$$

We can then write

$$(n-1) \frac{\mathbf{S}'_{\mathcal{K}}}{\mathbf{S}_{\mathcal{K}}} f' + f'' = knf.$$

Let $v = \mathbf{S}_{\mathcal{K}}(r)$ and $h(v) = f(\mathbf{S}_{\mathcal{K}}^{-1}(v))$; taking into account that

$$f'(r) = \mathbf{S}'_{\mathcal{K}}(r) h'(v) \quad \text{and} \quad f''(r) = \mathbf{S}''_{\mathcal{K}}(r) h'(v) + \mathbf{S}'_{\mathcal{K}}{}^2(r) h''(v),$$

we obtain

$$(n-1) \frac{\mathbf{S}'_{\mathcal{K}}{}^2(r) h'(v)}{\mathbf{S}_{\mathcal{K}}(r)} + \mathbf{S}''_{\mathcal{K}}(r) h'(v) + \mathbf{S}'_{\mathcal{K}}{}^2(r) h''(v) = knh(v).$$

Moreover, since

$$\mathbf{S}'_{\mathcal{K}}{}^2(r) + \mathcal{K} \mathbf{S}_{\mathcal{K}}^2(r) = 1 \quad \text{and} \quad \mathbf{S}''_{\mathcal{K}}(r) + \mathcal{K} \mathbf{S}_{\mathcal{K}}(r) = 0,$$

it turns out that

$$v(1 - \mathcal{K}v^2)h''(v) + (n(1 - \mathcal{K}v^2) - 1)h'(v) - knvh(v) = 0.$$

If we try $h(v) = \sum_{j=0}^{\infty} a_j v^j$, since

$$h'(v) = \sum_{j=1}^{\infty} a_j j v^{j-1}, \quad h''(v) = \sum_{j=2}^{\infty} a_j j(j-1) v^{j-2},$$

we have

$$\begin{aligned} \sum_{j=2}^{\infty} a_j j(j-1) v^{j-1} - \mathcal{K} \sum_{j=2}^{\infty} a_j j(j-1) v^{j+1} + (n-1) \sum_{j=1}^{\infty} a_j j v^{j-1} \\ - \mathcal{K} n \sum_{j=1}^{\infty} a_j j v^{j+1} - kn \sum_{j=0}^{\infty} a_j v^{j+1} = 0, \end{aligned}$$

that is

$$\begin{aligned} (n-1)a_1 + (2(n-1)a_2 + 2a_2 - kna_0)v \\ + \sum_{j=1}^{\infty} \{a_{j+2}(j+2)(n+j) - a_j(kn + \mathcal{K}j(n+j-1))\} v^{j+1} = 0. \end{aligned}$$

Now, if $n \neq 1$, then

$$a_1 = 0 \quad \text{and} \quad a_{j+2} = \frac{kn + \mathcal{K}j(n+j-1)}{(n+1)(j+2)} a_j, \quad j \geq 0.$$

Thus,

$$h(v) = a_0 \sum_{j=0}^{\infty} \frac{\prod_{s=1}^j \{kn + 2\mathcal{K}(s-1)(n+2s-3)\}}{j! \left(\frac{n}{2}\right)_j 4^j} v^{2j}$$

and

$$f(r) = a_0 \sum_{j=0}^{\infty} \frac{\prod_{s=1}^j \{kn + 2\mathcal{K}(s-1)(n+2s-3)\}}{j! \left(\frac{n}{2}\right)_j 4^j} \mathbf{S}_{\mathcal{K}}^{2j}(r),$$

where a_0 is determined by the condition $f(R) = -a$. It is easy to see that this series is convergent iff $|\mathcal{K}| \mathbf{S}_{\mathcal{K}}^2(r) < 1$. This is always true in the case of non-negative sectional curvatures.

Furthermore, we have to evaluate $\int_{S_R} f dV$. In spherical coordinates (see Appendix)

$$\int_{S_R} f dV = \text{ar}(S) \int_0^R \mathbf{S}_{\mathcal{K}}^{n-1} f(r) dr,$$

Since $(\mathbf{S}_{\mathcal{K}}^{n-1}(r)f'(r))' = kn\mathbf{S}_{\mathcal{K}}^{n-1}(r)f(r)$, we find that

$$\int_{S_R} f dV = \frac{2\pi^{n/2}}{kn\Gamma(n/2)} \mathbf{S}_{\mathcal{K}}^{n-1}(R)f'(R)$$

and

$$\mathcal{Y}^* = \frac{a^2}{kn} \text{vol}(S_R) \left(1 - \frac{f'(R)\mathbf{S}_{\mathcal{K}}^{n-1}(R)}{knf(R) \int_0^R \mathbf{S}_{\mathcal{K}}^{n-1}(r)dr} \right).$$

■

Corollary 2.3 . *When the parametric statistical model is an Euclidean manifold, we have the following lower bound for the Riemannian average of the mean squared Rao distance, on a ball of radius R :*

$$\mathcal{R}_{\mathcal{U}}^2(S_R) \geq \frac{n}{k} \left(1 - \frac{{}_0F_1\left(\frac{n}{2} + 1; \frac{knR^2}{4}\right)}{{}_0F_1\left(\frac{n}{2}, \frac{knR^2}{4}\right)} \right), \quad (9)$$

where ${}_0F_1(a; z)$ is a generalized hypergeometric function (see (16) of the Appendix).

If the Euclidean manifold M is complete and simply connected, we obtain the following lower bound in the manifold:

$$\mathcal{R}_{\mathcal{U}}^2(M) \equiv \lim_{R \rightarrow \infty} \mathcal{R}_{\mathcal{U}}^2(S_R) \geq \frac{n}{k}.$$

Proof: It is a particular case of the previous theorem with $\mathcal{K} = 0$. The second part of the proposition follows by taking the limit when $R \rightarrow \infty$ in (9). ■

Example 2.4 . As an example, consider the n -variate normal distribution with known covariance matrix Σ . Given a sample of size k , the Riemannian density of the mean squared Rao distance corresponding to the sample mean \bar{X}_k is $\mathcal{R}_{\mathcal{U}}^2(S_R) = n/k$, which coincides with the previous bound.

In the case $n = 1$, the manifold is Euclidean and we can apply the previous result; we obtain

$$\frac{\mathcal{Y}^*}{2R} = \frac{1}{k} \left(1 - \frac{\tanh(\sqrt{k}R)}{\sqrt{k}R} \right),$$

which coincides with the result already obtained by Čencov [3].

In fact, if we take a Cartesian coordinate system with origin p and try to solve the variational problem for a cube with center p , $C_R = \{x : |x_i| \leq R, i = 1, \dots, n\}$, we have to solve the Dirichlet problem:

$$\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} f = knf, \text{ with } f(x) = -n \text{ if } |x_i| = R \text{ for some } i = 1, \dots, n.$$

If we try $f(x) = \sum_{i=1}^n f_i(x_i)$, we obtain

$$\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} f_i(x_i) = k \sum_{i=1}^n f_i(x_i), \quad \sum_{i=1}^n f_i(\pm R) = -n.$$

Obviously, a solution is given by $f(x) = \sum_{i=1}^n g(x_i)$, with g such that

$$\frac{\partial^2}{\partial z^2} g(z) = kng(z), \quad g(\pm R) = -1.$$

The solution of the last equation is

$$g(z) = -\frac{\cosh \sqrt{kn}z}{\cosh \sqrt{kn}R},$$

and then

$$f(x) = -\sum_{i=1}^n \frac{\cosh \sqrt{kn}x_i}{\cosh \sqrt{kn}R}.$$

This provides the bound

$$\begin{aligned} \frac{\mathcal{Y}^*}{\text{vol}(C_R)} &= \frac{n}{k} \left(1 - \frac{g'(R)\text{ar}(C_R)}{kn^2g(R)\text{vol}(C_R)} \right) \\ &= \frac{n}{k} \left(1 - \frac{2nR\sqrt{kn} \tanh \sqrt{kn}R}{2kn^2R} \right) \\ &= \frac{n}{k} \left(1 - \frac{\tanh \sqrt{kn}R}{\sqrt{kn}R} \right), \end{aligned}$$

which improves upon the result given by Čencov ([3]). By Corollary (1.2), we can also give, in the general non-Euclidean case (fixed coordinate system), for the mean squared error (M.S.E.), a bound of the form

$$\mathcal{R}_U^2(C_R) \geq \frac{n^2}{k\eta} \left(1 - \frac{\tanh \sqrt{k\eta}R}{\sqrt{k\eta}R} \right),$$

where η is an upper bound of $\sum_{\beta} g_{\beta\beta}$ in C_R .

We can also give lower bounds to the general case

Theorem 2.5 . *When the parametric statistical model is a manifold with sectional curvatures bounded from above by \mathcal{K} , then we have the following lower bound for the average of the mean squared Rao distance:*

$$\mathcal{R}_U^2(S_R) \geq \frac{a^2}{kn} \left(1 - \frac{f'_{\mathcal{K}}(R)\text{ar}(S_R)}{knf_{\mathcal{K}}(R)\text{vol}(S_R)} \right) > 0, \quad (10)$$

where $\text{ar}(S_R)$ is the area of the n -dimensional sphere of radius R , $\text{vol}(S_R)$ its volume and $f_{\mathcal{K}}(r)$ is the solution of the boundary problem in S_R , (8), on a manifold of constant sectional curvature \mathcal{K} .

Proof: Consider a geodesic spherical coordinates (r, u) . Let $f(r, u)$ be the solution to the boundary problem (8) on a manifold with sectional curvatures bounded from above by \mathcal{K} . Let $f_{\mathcal{K}}(r)$ be the solution to the same problem but on a manifold of constant sectional curvature \mathcal{K} , which, as we know, only depends on the radial coordinate. Then,

$$\Delta f_{\mathcal{K}} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial r} \left(\sqrt{g} \frac{\partial}{\partial r} f_{\mathcal{K}} \right) = \frac{\partial^2}{\partial r^2} f_{\mathcal{K}} + \frac{\partial}{\partial r} (\log \sqrt{g}) \frac{\partial}{\partial r} f_{\mathcal{K}}.$$

By Bishop's comparison theorems we have

$$(n-1) \frac{S'_\kappa}{S_\kappa} \leq \frac{\partial}{\partial r} (\log \sqrt{g}),$$

and, since

$$\frac{\partial}{\partial r} f_\kappa \leq 0,$$

we have

$$\Delta f_\kappa - kn f_\kappa \leq \Delta_\kappa f_\kappa - kn f_\kappa = 0,$$

with Δ_κ the Laplacian for the constant sectional curvature case. Thus

$$\Delta f_\kappa - kn f_\kappa \leq \Delta f - kn f = 0.$$

Now, since $f_\kappa(p) = f(p) = -a$, $p \in \partial S_R$, we can apply the comparison theorem for elliptic differential equations (see Rauch, [9, Theorem 6 p. 243]). We find that

$$f(p) \leq f_\kappa(p), \quad p \in S_R,$$

and, since equality holds on the boundary,

$$\frac{\partial}{\partial r} f(p) \geq \frac{\partial}{\partial r} f_\kappa(p) \quad p \in \partial S_R.$$

Finally, by (4) and (1),

$$\mathcal{Y}^* = \frac{a^2}{kn} \text{vol}(S_R) + \frac{a}{kn} \int_{\partial S_R} \frac{\partial}{\partial r} f \, d\sigma \geq \frac{a^2}{kn} \text{vol}(S_R) \left(1 - \frac{f'_\kappa(R) \text{ar}(S_R)}{kn f_\kappa(R) \text{vol}(S_R)} \right),$$

and the proposition follows. ■

Remarks. Estimates for the volumes of balls given in the Appendix are useful to give a final expression for these bounds. Note that if the sectional curvatures are bounded from below by κ and from above by \mathcal{K} , by proposition (4.3), we have

$$\frac{\text{ar}(S_R)}{\text{vol}(S_R)} \leq \frac{S_\kappa^{n-1}(R)}{\int_0^R S_\mathcal{K}^{n-1}(r) dr}.$$

3 Lower bounds for the maximum risk

Even though we could use the Riemannian average of the risk to derive bounds for the maximum risk, we can obtain sharper minimax bounds and more directly

Lemma 3.1 *Let X be a smooth field on M such that $\operatorname{div}(X) \leq -a$, let f be a non-negative function on M and let W be a submanifold with boundary in M , then*

$$a \int_W f dV \leq \int_{\partial W} f \|X\| d\sigma + \int_W \|X\| \|\operatorname{grad}(f)\| dV$$

Proof:

$$\begin{aligned} a \int_W f dV &\leq \int_W (\langle X, \operatorname{grad}(f) \rangle - \operatorname{div}(fX)) dV \\ &\leq \int_W \|X\| \|\operatorname{grad}(f)\| dV + \int_{\partial W} f \|X\| d\sigma \end{aligned} \quad (11)$$

$$(12)$$

■

Theorem 3.2 . *We have the following lower bound for the local minimax risk of an estimator \mathcal{U} on W*

$$\sup_{p \in W} E(\rho^2(\mathcal{U}, p)) \geq \frac{a^2}{\left(\frac{\operatorname{ar}(\partial W)}{\operatorname{vol}(W)} + \sqrt{kn}\right)^2}$$

Proof: By the previous Lemma if we take $p = p_{(k)}(x; \theta)$, $X = \exp^{-1}(\mathcal{U})$, by integrating (11) with respect to $d\mu$ and by Fubini's theorem, we have

$$\begin{aligned} a \operatorname{vol}(W) &\leq \int_W \sqrt{E(\|A\|^2)} \sqrt{E(\|C\|^2)} dV + \int_{\partial W} E(\|A\|) d\sigma \\ &\leq \sqrt{kn} \int_W \sqrt{E(\|A\|^2)} dV + \int_{\partial W} \sqrt{E(\|A\|^2)} d\sigma. \end{aligned} \quad (13)$$

Thus

$$\sup_{p \in W} E(\|A\|^2) \geq \left(\frac{a \operatorname{vol}(W)}{\operatorname{ar}(\partial W) + \sqrt{kn} \operatorname{vol}(W)}\right)^2.$$

If the Euclidean manifold M is complete and simply connected, we obtain the following lower bound over the manifold:

$$\mathcal{R}_{\mathcal{U}}^2(M) \equiv \lim_{R \rightarrow \infty} \mathcal{R}_{\mathcal{U}}^2(S_R) \geq \frac{n(n+2)^2}{4k(n+1)^2}.$$

Proof: Since

$$\text{vol}(S_r) = \frac{2\pi^{n/2}r^n}{n\Gamma(n/2)},$$

we have

$$\int_0^R \sqrt{\text{vol}(S_r)} dr = \left(\frac{8\pi^{n/2}R^{n+2}}{n(n+2)^2\Gamma(n/2)} \right)^{1/2}$$

and

$$\int_0^R \text{vol}(S_r) dr = \frac{2\pi^{n/2}R^{n+1}}{n(n+1)\Gamma(n/2)};$$

then

$$0 < \left\{ \frac{n(n+2)R}{(n+1)(n+2+2\sqrt{kn}R)} \right\}^2 \leq \mathcal{R}_{\mathcal{U}_k}^2(S_R).$$

We derive the second statement taking the limit when $R \rightarrow \infty$. ■

Remarks. Example (2.4) shows that the bound obtained here is worse than the variational one if R goes to infinity, but it is better if R goes to zero:

$$\lim_{R \rightarrow 0} \left\{ \frac{\frac{n}{k} \left(1 - \frac{{}_0F_1\left(\frac{n}{2} + 1; \frac{knR^2}{4}\right)}{{}_0F_1\left(\frac{n}{2}; \frac{knR^2}{4}\right)} \right)}{\frac{n^2(n+2)^2R^2}{(n+1)^2(n+2+2\sqrt{kn}R)^2}} \right\} = \frac{(n+1)^2}{n(n+2)} > 1.$$

Remark. Note that global bounds for the average of the mean squared Rao distance also provide bounds for the local minimax risk in an obvious way. It can be shown these last bounds are sharper than bounds provided by the variational methods.

4 Appendix

4.1 Comparison theorems and volumes

We can use Bishop's theorems to obtain the volume of a ball of radius r in a Riemannian manifold whose sectional curvatures are constant and to give bounds for this volume when the sectional curvatures are bounded. We have the following propositions:

Proposition 4.1 . *If the sectional curvatures are constant and equal to \mathcal{K} , the volume of a Riemannian ball of radius r and center $p \in M$ is given by*

$$\text{vol}(S_r) = \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_0^r \mathbf{S}_{\mathcal{K}}^{n-1}(t) dt.$$

Proof: We have

$$\text{vol}(S_r) = \int_0^r \int_{\xi^{-1}(S_n)} \sqrt{g} du d\rho,$$

where S_n is the unit sphere in M_p . On the other hand, by Bishop's comparison theorems, when the sectional curvatures are constant,

$$\frac{\partial}{\partial \rho} \log \sqrt{g(\rho, u)} = (n-1) \frac{\mathbf{S}'_{\mathcal{K}}(\rho)}{\mathbf{S}_{\mathcal{K}}(\rho)},$$

with

$$\mathbf{S}_{\mathcal{K}}(t) = \begin{cases} \frac{\sin(\sqrt{\mathcal{K}}t)}{\sqrt{\mathcal{K}}} & \text{if } \mathcal{K} > 0, \\ t & \text{if } \mathcal{K} = 0, \\ \frac{\sinh(\sqrt{-\mathcal{K}}t)}{\sqrt{-\mathcal{K}}} & \text{if } \mathcal{K} < 0. \end{cases}$$

Then, integrating this expression, we have

$$\sqrt{g(\rho, u)} = \mathbf{S}_{\mathcal{K}}^{n-1} \Omega_{\mathcal{K}}(u).$$

However $\Omega_{\mathcal{K}}$ does not depend on \mathcal{K} . In fact

$$\lim_{\rho \rightarrow 0} \frac{\sqrt{g(\rho, u)}}{\rho^{n-1} \Omega(u)} = 1,$$

where $\Omega(u) du$ is the area element of the unit sphere in a Euclidean manifold and, since

$$\lim_{\rho \rightarrow 0} \frac{\mathbf{S}_\kappa}{\rho} = 1,$$

we conclude that $\Omega_\kappa = \Omega$. Thus, we may write

$$\text{vol}(S_r) = \int_{\xi^{-1}(S_n)} \Omega(u) du \int_0^r \mathbf{S}_\kappa^{n-1}(\rho) d\rho,$$

and finally,

$$\int_{\xi^{-1}(S_n)} \Omega(u) du = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

■

Proposition 4.2 . *When the sectional curvatures are constant and equal to \mathcal{K} and $|\mathcal{K}|\mathbf{S}_\kappa^2(r) < 1$, we have the following expression for the volume of a Riemannian ball of radius r :*

$$\text{vol}(S_r) = \frac{2\pi^{n/2}}{n \Gamma(n/2)} \mathbf{S}_\kappa^n(r) \left\{ 1 + \sum_{j=1}^{\infty} \frac{n\Gamma(j + \frac{1}{2})}{\sqrt{\pi}(n + 2j)} \frac{\mathcal{K}^j \mathbf{S}_\kappa^{2j}(r)}{j!} \right\}. \quad (15)$$

Proof: From the previous proposition,

$$\text{vol}(S_r) = \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_0^r \mathbf{S}_\kappa^{n-1}(t) dt.$$

Then, since by the definition of \mathbf{S}_κ ,

$$\{\mathbf{S}'_\kappa(t)\}^2 + \mathcal{K}\{\mathbf{S}_\kappa(t)\}^2 = 1,$$

and, making the change $y = \mathbf{S}_\kappa^2(t)/\mathbf{S}_\kappa^2(r)$, we have

$$\int_0^r \mathbf{S}_\kappa^{n-1}(t) dt = \frac{1}{2} \mathbf{S}_\kappa^n(r) \int_0^1 y^{\frac{n-2}{2}} (1 - \mathcal{K} \mathbf{S}_\kappa^2(r) y)^{-\frac{1}{2}} dy.$$

Moreover, there is a relationship between integrals of this kind and the generalized hypergeometric functions. These functions are defined by,

$${}_pF_q(a_1, \dots, p; b_1, \dots, b_q; z) \equiv \sum_{j=0}^{\infty} \frac{(a_1)_j \cdots (a_p)_j z^j}{(b_1)_j \cdots (b_q)_j j!}, \quad (16)$$

where $(a)_j = a(a+1)\cdots(a+j-1)$ and z is any complex number if $p \leq q$, $\|z\| < 1$ if $p = q + 1$ and they diverge for all $z \neq 0$ if $p > q + 1$ (see Abramowitz [1]). The relationship is the following:

$${}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-tz)^{-a} dt, \quad \operatorname{Re}(c) > \operatorname{Re}(b) > 0.$$

This leads to

$$\begin{aligned} \int_0^r \mathbf{S}_{\mathcal{K}}^{n-1}(t) dt &= \frac{1}{2} \mathbf{S}_{\mathcal{K}}^n(r) \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+2}{2})} {}_2F_1\left(\frac{1}{2}, \frac{n}{2}; \frac{n+2}{2}; \mathcal{K} \mathbf{S}_{\mathcal{K}}^2(r)\right) = \\ &= \frac{1}{2} \mathbf{S}_{\mathcal{K}}^n(r) \frac{1}{\Gamma(\frac{1}{2})} \sum_{j=0}^{\infty} \frac{\Gamma(\frac{1}{2} + j)}{\frac{n}{2} + j} \frac{\mathcal{K}^j \mathbf{S}_{\mathcal{K}}^{2j}(r)}{j!}, \end{aligned}$$

and the proposition is proved. ■

Proposition 4.3 . *Let $\operatorname{vol}(S_r(p))$ be the volume of a ball with center p and radius r , on a manifold with sectional curvatures bounded from below by κ and from above by \mathcal{K} . Then*

$$\operatorname{vol}_{\kappa}(S_r) \geq \operatorname{vol}(S_r(p)) \geq \operatorname{vol}_{\mathcal{K}}(S_r),$$

where $\operatorname{vol}_{\kappa}(S_r)$ and $\operatorname{vol}_{\mathcal{K}}(S_r)$ are, respectively, the volumes of balls of radius r and arbitrary centers on manifolds with constant sectional curvatures κ and \mathcal{K} .

Proof: If we integrate, from ρ_0 to ρ , the inequalities in Bishop's comparison theorems, we obtain

$$\frac{\mathbf{S}_{\kappa}^{n-1}(\rho)}{\mathbf{S}_{\kappa}^{n-1}(\rho_0)} \geq \frac{\sqrt{g(\rho, u)}}{\sqrt{g(\rho_0, u)}} \geq \frac{\mathbf{S}_{\mathcal{K}}^{n-1}(\rho)}{\mathbf{S}_{\mathcal{K}}^{n-1}(\rho_0)}.$$

Moreover,

$$\lim_{\rho_0 \rightarrow 0} \sqrt{g(\rho_0, u)} \frac{\mathbf{S}_{\kappa}^{n-1}(\rho)}{\mathbf{S}_{\kappa}^{n-1}(\rho_0)} \geq \sqrt{g(\rho, u)} \geq \lim_{\rho_0 \rightarrow 0} \sqrt{g(\rho_0, u)} \frac{\mathbf{S}_{\mathcal{K}}^{n-1}(\rho)}{\mathbf{S}_{\mathcal{K}}^{n-1}(\rho_0)},$$

and, since

$$\lim_{\rho_0 \rightarrow 0} \frac{\sqrt{g(\rho_0, u)}}{\mathbf{S}_\kappa^{n-1}(\rho_0)} = \lim_{\rho_0 \rightarrow 0} \frac{\sqrt{g(\rho_0, u)}}{\mathbf{S}_\kappa^{n-1}(\rho_0)} = \Omega(u),$$

we conclude that

$$\mathbf{S}_\kappa^{n-1}(\rho)\Omega(u) \geq \sqrt{g(\rho, u)} \geq \mathbf{S}_\kappa^{n-1}(\rho)\Omega(u).$$

■

References

- [1] M. Abramowitz. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1970.
- [2] J. O. Berger. *Statistical decision theory and Bayesian analysis*. Springer series in statistics, New York, 1985.
- [3] N.N. Cencov. *Statistical Decision Rules and Optimal Inference*. Nauka, Moscow (en ruso, versi n inglesa, 1982, Math. Monographs, 53, Amer. Math. Soc., Providence), 1972.
- [4] I. Chavel. *Eigenvalues in Riemannian Geometry*. Pure and applied mathematics. Academic Press, Inc., Orlando, Florida, 1984.
- [5] N.J. Hicks. *Notes on Differential Geometry*. Van Nostrand Reinhold Comp., London, 1965.
- [6] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. Royal Soc.*, 186 A:453–461, 1946.
- [7] J.M. Oller and J.M. Corcuera. Intrinsic analysis of statistical estimation. *To appear in The Annals of Statistics*, 1993.
- [8] B.L.S. Prakasa Rao. Cram r-rao type integral inequalities for estimators of functions of multidimensional parameter. *Sankhyà*, 54 A:53–73, 1992.
- [9] J. Rauch. *Partial Differential Equations*. Springer–Verlag, New York, 1991.

Relació dels últims Preprints publicats:

- 194 *Some compactness results in real interpolation for families of Banach spaces.* María J. Carro and Jaak Peetre. 1980 Mathematics Subject Classification (1985 Revision): 46M37, 47B07. January 1996.
- 195 *Separatrix splitting for a high-frequency perturbation of the pendulum.* V. G. Gelfreich. 1991 Mathematics Subject Classification: 34C37, 34C29, 58F30, 70H05. January 1996.
- 196 *A second order Stratonovich differential equation with boundary conditions.* Aureli Alabert and David Nualart. AMS Subject Classification: 60H10, 60H07. January 1996.
- 197 *A Gentzen system equivalent to the BCK-logic.* Romà J. Adillon and Ventura Verdú. Mathematics Subject Classification: 03B20. January 1996.
- 198 *On Gentzen systems associated with the finite linear MV-algebras.* Àngel J. Gil, Antoni Torrens Torrell and Ventura Verdú. Mathematics Subject Classification: 03B22, 03F03, 03G20, 03G99. February 1996.
- 199 *Exponentially small splitting of separatrices under fast quasiperiodic forcing.* Amadeu Delshams, Vassili Gelfreich, Àngel Jorba and Tere M. Seara. AMS classification scheme numbers: 34C37, 58F27, 58F36, 11J25. February 1996.
- 200 *Existence and regularity of the density for solutions to stochastic differential equations with boundary conditions.* Arturo Kohatsu-Higa and Marta Sanz-Solé. AMS 1990 SCI: 60H07, 60H10, 60H99. March 1996.
- 201 *A forcing construction of thin-tall Boolean algebras.* Juan Carlos Martínez. 1991 Mathematics Subject Classification: 03E35, 06E99, 54612. March 1996.
- 202 *Weighted continuous metric scaling.* C. M. Cuadras and J. Fortiana. AMS Subject Classification: 62H25, 62G99. April 1996.
- 203 *Homoclinic orbits in the complex domain.* V.F. Lazutkin and C. Simó. AMS Subject Classification: 58F05. May 1996.
- 204 *Quasivarieties generated by simple MV-algebras.* Joan Gispert and Antoni Torrens. Mathematics Subject Classification: 03B50, 03G99, 06F99, 08C15. May 1996.
- 205 *Regularity of the law for a class of anticipating stochastic differential equations.* Carles Rovira and Marta Sanz-Solé. AMS Subject Classification: 60H07, 60H10. May 1996.
- 206 *Effective computations in Hamiltonian dynamics.* Carles Simó. AMS Subject Classification: 58F05, 70F07. May 1996.
- 207 *Small perturbations in a hyperbolic stochastic partial differential equation.* David Márquez-Carreras and Marta Sanz-Solé. AMS Subject Classification: 60H15, 60H07. May 1996.
- 208 *Transformed empirical processes and modified Kolmogorov-Smirnov tests for multivariate distributions.* A. Cabaña and E. M. Cabaña. AMS Subject Classification: Primary: 62G30, 62G20, 62G10. Secondary: 60G15. June 1996.
- 209 *Anticipating stochastic Volterra equations.* Elisa Alòs and David Nualart. AMS Subject Classification: 60H07, 60H20. June 1996.
- 210 *On the relationship between α -connections and the asymptotic properties of predictive distributions.* J.M. Corcuera and F. Giummolè. AMS 1980 Subjects Classifications: 62F10, 62B10, 62A99. July 1996.



