# Shrinkage estimators under spherically symmetry for the general linear model with respect to general quadratic loss

by

Dominique Cellier and Dominique Fourdrinier
Université de Rouen, France

BU-1157-M

May 1992

**Abstract**

This paper is primarily concerned with extending the results of Brandwein and Strawderman [4] and [5] in the usual canonical setting of a general linear model when sampling from a spherically symmetric distribution. When the location parameter belongs to a proper linear subspace of the sampling space, we give an unbiased estimator of the difference of the risks between the least squares estimator $\varphi_0$ and a general shrinkage estimator $\varphi = \varphi_0 - \| X - \varphi_0 \|^2 \cdot g o \varphi_0$. We obtain a general condition of domination for $\varphi$ over $\varphi_0$ which is weaker than that of Brandwein and Strawderman. We do not need superharmonicity condition on g. Our results are valid for general quadratic loss.

AMS 1980 subject classifications: 62C99, 62F10, 62H99.

*Key words and phrases:* spherical symmetry, minimaxity, quadratic loss, James-Stein estimation, location parameter, superharmonic functions.

## 1. Introduction

This paper is concerned with estimating a location vector in the framework of the general linear model. In order to underline the intrinsic aspect of our results, the approach of multivariate analysis adopted here is "coordinate free" (see Kruskall [11] and [12] and Stone [18]). The set-up agrees with the article of Cellier, Fourdrinier and Robert [8].

The observation x belongs to an n-dimensional euclidian space E and the location parameter to a proper k-dimensional linear subspace $\Theta$ ($0 < k < n$). The underlying distributions $P_\theta$ are assumed to be spherically symmetric around $\theta$ (see section 2). The loss functions are general quadratic loss functions.

It is worth noting that we do not assume that $P_\theta$ has a density with respect to the Legesgue measure on E.

As $k < n$, the usual estimator of $\Theta$ is the orthogonal projector $\varphi_0$ from E onto $\Theta$ (the least square estimator, which is minimax) and the competing estimators considered are of the form $\varphi = \varphi_0 - \| X - \varphi_0 \|^2 \cdot g o \varphi_0$ where X is the identity function in E (for every $x \in E$, $X(x) = x$) and g is a measurable function from $\Theta$ into $\Theta$.

This paper is concerned with extending the results of Brandwein and Strawderman [4] and [5], particularly the results in the general set-up of the linear model in [4] (section 5 and examples of the section 2). This is generalized in [5] for a general quadratic loss (section 3). Specifically we show that the estimators $\varphi$ dominates $\varphi_0$ under weaker conditions than those they use, namely $\| g \|^2 \le \frac{2}{n-k+2} \operatorname{div} g$. We do not need either superharmonicity conditions on a certain function h such that $\frac{\| g \|^2}{2} \le h \le \operatorname{div} g$, or assumption of nonincreasing for the function $R \to R^2 E[h]$ where the expectation is considered with respect to the uniform distribution on the ball with radius R and center $\theta$. This later result is important for robustness.

The difference between the risk function of $\varphi_0$ and the risk function of $\varphi$ is given by

$$\delta(\theta) = 2\, E_\theta\left[\| X - \varphi_0 \|^2 < \varphi_0 - \theta, h o \varphi_0 >\right] - E_\theta\left[\| X - \varphi_0 \|^4 \| h o \varphi_0 \|^2\right] \tag{1.1}$$

where $E_\theta$ denotes the expectation with respect to $P_\theta$. Hence the new estimator $\varphi$ is at least as good as the usual one $\varphi_0$ if $\delta(\theta) \geq 0$.

Since $P_\theta$ is spherically symmetric, for every bounded random variable f, we have

$$E_\theta[f] = \int_R E_{R,\theta}\,[f]\;\rho(dR)$$

where $E_{R,\theta}$ is the expectation with respect to the uniform distribution, denoted by $U_{R,\theta}$, on the sphere $S_{R,\theta} = \{x \in E : || x || = R\}$ of radius R and center $\theta$, and $\rho$ is the distribution of the radius, namely the distribution of the norm $||\ ||$ under $P_0$. It suffices to prove the result working conditionally on the radius, that is to say to replace $P_\theta$ by $U_{R,\theta}$ in the expression (1.1). Then, for a given R, noticing that for every $x \in S_{R,\theta}$,

$$|| x - \varphi_0(x) ||^2 = R^2 - || \varphi_0(x) - \theta ||^2,$$

we have to consider

$$\delta_R(\theta) = 2\ E_{R,\theta}\Big[(R^2 - || \varphi_0 - \theta ||^2) < \varphi_0 - \theta,\ ho\varphi_0 >\Big] - E_{R,\theta}\Big[\big(R^2 - || \varphi_0 - \theta ||^2\big)^2 || ho\varphi_0 ||^2\Big]\ . \qquad (1.2)$$

Since the integrand terms in these integrals depend on the observation only through $\varphi_0$, (1.2) can be calculated using the fact that the distribution of $\varphi_0$, under $U_{R,\theta}$, has a density with respect to the Lebesgue measure on $\Theta$ (see lemma in section 5).

Brandwein and Strawderman [4] and [5] also use conditioning, however they work conditionally on $|| \varphi_0 || = r$ and $|| X - \varphi_0 || = s$. The argument in [4] under the usual quadratic loss proceeds as follows.

Using the divergence theorem for the cross-product term (i.e., the first term in (1.1)), they obtain

$$2\,\frac{s^2 r^2}{k} \int_{B_{r,\theta}} \mathrm{div}\ g(t)\ V_{r,\theta}(dt) - s^4 \int_{S_{r,\theta}} || g(t) ||^2\ U_{r,\theta}(dt)\ ,$$

where $V_{r,\theta}$ is the uniform distribution on the ball $B_{r,\theta} = \{x \in E : || x - \theta || \leq r\}$. Afterward, in order to compare these two integrals, they need an extra assumption that there exists a superharmonic function h such that

$$\frac{\| g \|^2}{2} \leq h \leq \text{div } g \ .$$

So, using the evident notation $\delta_{r,s}(\theta)$ for the difference between the risks conditionally on r and s, they obtain the lower bound for $\delta_{r,s}(\theta)$

$$\delta_{r,s}(\theta) \geq 2 \, \frac{s^2 r^2}{k} \int_{B_{r,\theta}} h(t) \, V_{r,\theta}(dt) - 2s^4 \int_{B_{r,\theta}} h(t) \, V_{r,\theta}(dt) \, . \tag{1.3}$$

Moreover, since the two terms in the right member of (1.3) are not homogeneous in s and r they need another extra assumption about h, that is, the function $r \to r^2 E_{V_{r,\theta}}[h]$ is nondecreasing.

Our method is more powerful in the case where the dimension of the parameter space $\Theta$ is less than the dimension of the sample space E. Indeed we also use conditioning, however we do not work conditionally on $\| \varphi_0 \|^2 = r$ and $\| X - \varphi_0 \|^2 = s$ but conditionally on the radius, $\| X - \theta \| = R$. Then we use the fact that, although $U_{R,\theta}$ is a singular distribution, its image by the projector $\varphi_0$ is absolutely continuous with respect to the Legesgue measure on $\Theta$. Thus the corresponding density seems to be a good tool, since it permits one to get an exact expression of the risk difference and not only a lower bound. Its use exploits the information well that $\theta$ belongs to a proper linear subspace.

Section 2 provides the model and the main result. In Section 3 we consider a general example. Section 4 states the consequences of the finiteness of the risk of the shrinkage estimator. The lemma which gives the density of $\varphi_0$ under $U_{R,\theta}$ constitutes Section 5.

## 2. Shrinkage estimators with respect to general quadratic loss

### 2.1 Model

Let (E, <, >) be an n-dimensional euclidian space (<,> denoting the inner product and $\| \ \|$ denoting the norm which is connected with it) and $\Theta$ a proper k-dimensional linear subspace of E $(0 < k < n)$.

Let x be an observation of a spherically symmetric distribution $P_\theta$ around a vector $\theta$. Recall that spherically symmetry is equivalent to the fact that $P_\theta$ is the image of a distribution which stays invariant under any orthogonal transformation (with respect to the inner product <,>) translated by $\theta$.

The main assumption in this model (which coincides with the canonical form of the general linear model, cf.[15]) is that the location parameter $\theta$ belongs to $\Theta$ with dim $\Theta$ less than dim E.

The location parameter $\theta$ is estimated by measurable functions $\varphi$ from E into $\Theta$. In order to compare different estimators we consider a general quadratic loss connected with a positive quadratic form on $\Theta$. The loss incurred by the estimator $\varphi(x)$ when $\theta$ is the true value of the parameter is $q(\varphi(x)-\theta)$. Thus an estimator $\varphi$ improves an estimator $\varphi'$ if the corresponding risk $R(\varphi, .)$ of $\varphi$ is less than or equal to the risk $R(\varphi', .)$ of $\varphi'$, i.e. if, for every $\theta \in \Theta$,

$$R(\varphi,\theta) = \int_E q(\varphi(x)-\theta)\, P_\theta(dx) \le \int_E q(\varphi'(x)-\theta)\, P_\theta(dx) = R(\varphi',\theta) \ .$$

The goal of this paper is to give general conditions for a wide class of shrinkage estimators which improve the usual least square estimator. The shrinkage estimators we consider are of the form $\varphi(x) = \varphi_0(x) - \| x - \varphi_0(x) \|^2 \cdot g\big(\varphi_0(x)\big)$. We will denote this as

$$\varphi = \varphi_0 - \| X - \varphi_0 \|^2 \cdot go\varphi_0 \tag{2.1}$$

where g is a measurable application from $\Theta$ into $\Theta$.

**Remark 2.1.1.**

It is important to note that the shrinkage function $\| X - \varphi_0 \|^2 \cdot go\varphi_0$ includes the "residual term" $\| X - \varphi_0 \|^2$. It has been observed by Cellier, Fourdrinier and Robert (c.f. [3 ]) that the inclusion of this term yields some robustness properties while its absence yields nonrobustness. Since, for a given observation x, the residual term $\| x - \varphi_0(x) \|^2$ represents the square of the distance between x and its projection on $\Theta$, it is intuitively clear that its consideration improves the information use of the estimator.

## 2.2 Conditions on the distributions and the shrinkage function

In order to assure the finiteness of the risk of the usual estimator $\varphi_0$ and the risk of the shrinkage estimator $\varphi$ we need the two hypotheses (H1) and (H2), which are

(H1) $$\int_E \| \varphi_0 \|^2 \, dP_0 < +\infty$$

(H2) $$\forall \theta \in \Theta \quad \int_E \| X - \varphi_0 \|^4 \cdot \| g o \varphi_0 \|^2 \, dP_\theta < +\infty \, .$$

**Remark 2.2.1**

At first sight, the hypothesis (H1) is weaker than the fact that $P_0$ has a finite second moment, i.e. $E_0[\| X \|^2] < +\infty$. However, it is not the case. Indeed it is easy to show that, since $P_0$ is spherically symmetric around 0, these two hypotheses are equivalent.

**Remark 2.2.2**

It will often be difficult to check the condition (H2) by itself. On the other hand we show, in section 4, that the following condition

$$\text{(H3)} \quad \exists \, \alpha > 0 \quad \forall \, t \in \Theta \quad \| t \|^2 \; \| g(t) \|^2 \le \alpha$$

is sufficient to imply (H2). Such a type of condition is always needed in the literature (see, for instance, Berger [1] and [2], Bock [3], Cellier, Fourdrinier and Robert [8]), appears close to the original one (H2) (see section 4 for more details) and seems a generalization, when the shrinkage factor is vectorial, of those usually met.

A last important point is that we show, in section 4, the condition (H3) imposes that the dimension of $\Theta$ is greater than 2. Thus we find again the dimension condition for the admissibility of the least squares estimator, first obtained by Stein [16] in the normal case and generalized by Brown [6] for more general distributions.

We shall assume $k \ge 3$ in the following.

**2.3 Main Result**

Assuming the hypotheses (H1) and (H2), the differentiability of the shrinkage factor g and $k \ge 3$, our main result about the shrinkage estimator $\varphi$ of the form (2.1) is given in theorem 2.3.1 and will come from the following proposition 2.3.1.

**Proposition 2.3.1**

An unbiased estimator of the difference of the risks between $\varphi_0$ and $\varphi$ is

$$\left(\frac{2}{n-k+2}\,\mathrm{div}(Q o g o \varphi_0) - q o g o \varphi_0\right) \| X - \varphi_0 \|^4 \;.$$

A necessary and sufficient condition for domination of $\varphi_0$ by $\varphi$ is

$$\inf_{\theta \in \Theta} \frac{\int_E \mathrm{div}(Q o g o \varphi_0) \| X - \varphi_0 \|^4 \, dP_\theta}{\int_E q o g o \varphi_0 \| X - \varphi_0 \|^4 \, dP_\theta} \geq \frac{n-k+2}{2} \,.$$

where Q is the endomorphism connected with the quadratic form q and the inner product <,>.

**Remark 2.3.1**

Precisely the endomorphism Q, the inner product < , > and the quadratic form q are connected with the relation

$$\forall\, \theta \in \Theta \qquad q(\theta) = <\theta, Q(\theta)> .$$

If we were not in the context of the coordinate free approach, we would use the terminology of matrix theory. However when dealing with the coordinate free approach we no longer need to rely on the existence of a basis, hence we apply the notion of an endomorphism in our quadratic forms.

<u>Proof</u>

First let us notice that, since the ratio of two positive quadratic forms is bounded, the conditions (H1) and (H2) are respectively equivalent to (H1)′ and (H2)′ where

$$(\mathrm{H1})' \quad \int_E q(\varphi_0) dP_0 < +\infty$$

$$(\mathrm{H2})' \quad \forall\, \theta \in \Theta \quad \int_E \| x - \varphi_0 \|^4 \cdot q o g o \varphi_0 dP_\theta \,.$$

Let us fix $\theta$ in $\Theta$. According to condition (H1)′ and the linearity of $\varphi_0$, $R(\varphi_0, \theta)$ is finite since

$$R(\varphi_0, \theta) = \int_E q(\varphi_0(x) - \theta)\, P_\theta(dx) = \int_E q(\varphi_0(x-\theta))\, P_\theta(dx) = \int_E q(\varphi_0(x)\, P_0(dx) < +\infty.$$

Considering the risk difference

$$\delta(\theta) = R(\varphi_0, \theta) - R(\varphi, \theta)$$

between the risk of $\varphi_0$ and the risk of $\varphi$ at $\theta$, the finiteness of $R(\varphi, \theta)$ result from the condition (H2)$'$. Indeed we have

$$\delta(\theta) = 2\int_E \| X - \varphi_0 \|^2 < \varphi_0 - \theta, Qog o\varphi_0 > dP_\theta - \int_E \| X - \varphi_0 \|^4 \, qogo\varphi_0 \, dP_\theta < +\infty.$$

The hypothesis (H2)$'$ assures the second term of the right member of this equality is finite and the Schwarz inequality assures the first term is also finite.

We can now calculate the expression of the difference of the risks $\delta(\theta)$. Since $P_\theta$ is spherically symmetric around $\theta$, we have

$$\delta(\theta) = \int_{\mathbb{R}_+} \delta_R(\theta)\rho(dR)$$

with

$$\delta_R(\theta) = 2\int_{S_{R,\theta}} \| X - \varphi_0 \|^2 < \varphi_0 - \theta, Qogo\varphi_0 > dU_{R,\theta} - \int_{S_{R,\theta}} \| X - \varphi_0 \|^4 \, qogo\varphi_0 \, dU_{R,\theta}$$

where the notations $U_{R,\theta}$, $S_{R,\theta}$ and $\rho$ were introduced in Section 1.

Thus the results given in the statement of the proposition will follow, by working conditionally on the radius, that is with $\delta_R(\theta)$. Recall that, for every $x \in S_{R,\theta}$,

$$\| (X - \varphi_0)(x) \|^2 = R^2 - \| \varphi_0 (x) - \theta \|^2 .$$

Therefore the integrand term in the two integrals occuring in the expression of $\delta_R(\theta)$ depends only on $\varphi_0$ and thus using the density of $\varphi_0$ under $U_{R,\theta}$, we obtain (the notation $B_{R,\theta}$ being introduced in Section 1)

$$\delta_R(\theta) = 2C_R^{n,k}\int_{B_{R,\theta}} < t-\theta, (Qog)(t) > (R^2 - \|t-\theta\|^2)^{\frac{n-k}{2}} dt - C_R^{n,k}\int_{B_{R,\theta}} (qog)(t)(R^2 - \|t-\theta\|^2)^{\frac{n-k}{2}+1} dt \quad (2.3)$$

where $C_R^{n,k}$ is the normalization coefficient of the density (see lemma in section 5).

The result comes from the handling of the cross-product term (i.e. the first integral in the right member of (2.3)). The key fact is that the vector $(R^2 - \| t-\theta \|^2)^{\frac{n-k}{2}} (t-\theta)$ may be represented as the gradient at the point t of the function

$$\gamma : t \to \frac{-(R^2 - \| t-\theta \|^2)^{\frac{n-k}{2}+1}}{n-k+2} .$$

Thus, using the identity

$$\forall t \in \Theta \ \ \mathrm{div}(\gamma \cdot Qog)(t) = < \nabla \gamma(t), (Qog)(t) > + \gamma(t) \ \ \mathrm{div}\ (Qog)(t)$$

we can write

$$\int_{B_{R,\theta}} <t-\theta, (Qog)(t)>(R^2 - \| t-\theta \|^2)^{\frac{n-k}{2}} dt = \int_{B_{R,\theta}} \mathrm{div}(\gamma \cdot Qog)(t)dt - \int_{B_{R,\theta}} \gamma(t)\mathrm{div}(Qog)(t)dt. \tag{2.4}$$

Now the divergence theorem permits us to write the first integral of the right member of (2.4) as

$$\int_{B_{R,\theta}} \mathrm{div}(\gamma \cdot Qog)(t)dt = \int_{S_{R,\theta}} < (\gamma \cdot Qog)(t), \frac{t-\theta}{\| t-\theta \|} > \sigma_{R,\theta}\,(dt) \tag{2.5}$$

where $\sigma_{R,\theta}$ is the area measure on the sphere $S_{R,\theta}$ of radius R and center $\theta$ in $\Theta$, and assures that this term is null since the function $\gamma$ is identically equal to zero on $S_{R,\theta}$. It follows from (2.3), (2.4) and (2.5) that

$$\delta_R(\theta) = C_R^{n,k} \int_{B_{R,\theta}} \Big(\frac{2}{n-k+2}\, \mathrm{div}(Qog)(t) - (qog)(t)\Big)\Big(R^2 - \| t-\theta \|^2\Big)^{\frac{n-k}{2}+1} dt \ .$$

Hence, coming back to the expectation with respect to $P_\theta$,

$$\delta(\theta) = \int_E \Big(\frac{2}{n-k+2}\, \mathrm{div}(Qogo\varphi_0) - qogo\varphi_0\Big) \| X - \varphi_0 \|^4 \, dP_\theta \, .$$

Thus this integral means that $\Big(\frac{2}{n-k+2}\, \mathrm{div}(Qogo\varphi_0) - qogo\varphi_0\Big) \| X - \varphi_0 \|^4$ is an unbiased estimator of the difference of the risks between $\varphi_0$ and $\varphi$. The necessary and sufficient condition of domination of $\varphi$ over $\varphi_0$ also follow directly from this expression. □

Now we can state our main result which is an immediate consequence of the proposition.

**Theorem 2.3.1**

Assume that the hypotheses (H1) and (H2) are satisfied, the shrinkage factor g is differentiable and $k \geq 3$.

A sufficient condition for domination of $\varphi_0$ over $\varphi$ is

$$q \circ g \leq \frac{2}{n-k+2} \operatorname{div} (Q \circ g) .$$

**Remark 2.3.2**

As C. Stein [17], in the normal case, we obtain an unbiased estimator of the difference of the risk. This is due to the use of the density of $\varphi_0$, which appears as a very powerful tool. This type of result cannot be obtained by Brandwein and Strawderman [4] and [5] since they use a lower bound for the difference of the risks in order to prove the domination of $\varphi$.

Of course we find again the important property of robustness already noticed by Cellier, Fourdrinier and Robert [8], since by Brandwein and Strawderman [4].

A last point which is interesting to notice is the fact that the proof of the theorem does not need the use of a density for $P_\theta$.

## 3. A general example

A general example of nonspherical shrinkage estimator is given by the choice of a shrinkage factor g defined, for every $t \in \Theta$, by

$$g(t) = r(\| t \|^2) \frac{A(t)}{b(t)}$$

where r is a positive differentiable and nondecreasing function, A is a symmetric endomorphism whose eigenvalues are positive and b is a positive definite quadratic form on $\Theta$.

For every endomorphism C on $\Theta$, we denote by $C_M$, $C_m$ and tr(C) the maximum eigenvalue, the minimum eigenvalue and the trace of C. Likewise, if d is a quadratic form on $\Theta$, we use the same notations for d as those for the endomorphism connected with d and the inner product < , >.

Since, for every $t \in \Theta$,

$$\| t \| \; \| g(t) \| = r(\| t \|^2) \frac{\| t \|}{b(t)} \| A(t) \| \leq \frac{A_M}{b_m} r(\| t \|^2)$$

the condition (H3) is satisfied provided that the function r is bounded; thus the risk of the shrinkage estimator is finite (see section 4).

For every $t \in \Theta$, we have

$$(qog)(t) = \frac{r^2(\| t \|^2)}{b^2(t)} \cdot (qoA)(t) \leq \frac{(Q^{1/2}oA)_M^2}{b_m} \frac{r^2(\| t \|^2)}{b(t)}$$

where $Q^{1/2}$ is the endomorphism satisfying $Q = Q^{1/2}oQ^{1/2}$, and

$$\mathrm{div}(Qog)(t) = \left\langle \nabla \frac{1}{b}(t), r(\| t \|^2),(QoA)(t) \right\rangle + \frac{1}{b(t)} \left( \left\langle \nabla r(\| t \|^2),(QoA)(t) \right\rangle + r(\| t \|^2)\mathrm{div}(QoA)(t) \right)$$

$$= - \frac{r(\| t \|^2)}{b^2(t)} \left\langle \nabla b(t),(QoA)(t) \right\rangle + \frac{1}{b(t)} \left( 2r'(\| t \|^2) \left\langle t,(QoA)(t) \right\rangle + \mathrm{tr}(QoA)\, r(\| t \|^2) \right)$$

$$\geq - 2(QoA)_M \frac{r(\| t \|^2)}{b(t)} + 2\frac{(QoA)_m}{b_M} r'(\| t \|^2) + \mathrm{tr}(QoA) \frac{r(\| t \|^2)}{b(t)} ,$$

that last inequality being obtained considering usual lower and upper bounds about the ratios of positive quadratic forms and the fact that the functions r and r′ are nonnegative. Then it is clear that, in order that the condition (2.2) is satisfied, it suffices that, for every $t \in \Theta$,

$$\frac{(Q^{1/2}oA)_M^2}{b_m} \frac{r^2(\| t \|^2)}{b(t)} \leq \frac{2}{n-k+2} \left( - 2(QoA)_M \frac{r(\| t \|^2)}{b(t)} + 2\frac{(QoA)_m}{b_M} r'(\| t \|^2) + \mathrm{tr}(QoA) \frac{r(\| t \|^2)}{b(t)} \right)$$

which is equivalent to

$$\left( \frac{(Q^{1/2}oA)_M^2}{b_m} r(\| t \|^2) + \frac{2\left(2(QoA)_M - \mathrm{tr}(Qoa)\right)}{n-k+2} \right) \frac{r(\| t \|^2)}{b(t)} \leq \frac{4(QoA)_m}{(n-k+2)b_M} r'(\| t \|^2) \; .$$

Since the function $r'$ is non-negative, this condition is satisfied provided that, for every $t \in \Theta$,

$$r(\| t \|^2) \leq 2\frac{tr(QoA) - 2(QoA)_M}{n-k+2} \cdot \frac{b_m}{(Q^{1/2}oA)_M^2} .$$

**Remark 3.1**

The example above contains the examples 2.1 and 2.2 of Brandwein and Strawderman [4] and their example 3.1 in [5]. It is worth noting that the result of domination imposes only $k \geq 3$ (since we must have $tr(Qoa) - 2(QoA)_M > 0$), when the Brandwein and Strawderman result needs $k \geq 4$. It is explained by the fact that, when $k = n$, their method needs that div g is superharmonic which is true for $k \geq 4$. For instance, this requirement is easy to see when the function r is constant, the endomorphism A is the identity on $\Theta$ and the quadratic form b is the square of the usual norm (so $\varphi$ is the James-Stein estimator), since in this case $\Delta\left(\text{div } g(t)\right) = \frac{-r\,(k-2)(k-4)}{\| t \|^4}$.

When we compare their examples (in the framework of their theorem 3.1 in [5]), it is easy to see that we get a larger bound for r. Indeed their bound is equal to our bound generally multiplied by $\frac{k-2}{k}$ and particularly multiplied by $\frac{k-2}{k} \cdot \frac{b_m}{b_M}$ for the second choice of function h (see section 1 for the usefulness of function) in the example 2.2 of [4].

At least we do not use superharmonicity assumption. Thus we do not need the concavity of the function r which is required by Brandwein and Strawderman [4].

## 4. About the finiteness of the risk of $\varphi$

We noticed in section 2 that the finiteness of the risk of $\varphi$ is assured by the hypothesis (H2). This question is often implicit in the works about shrinkage estimators (see [4] and [5]) but deserves to be studied. The following gives some hints which indicate that the condition (H2) is not such a strong condition, especially since it does not give rise to a reduction of the class of the spherically symmetric distributions (which is the class of the spherically symmetric distributions with finite second moment according to (H1) and remark 2.2.1); so it is mainly a condition on the shrinkage factor g. Actually we can easily show that using the stronger condition (H3).

The condition (H2) states, for $\theta$ fixed in $\Theta$, that

$$B = \int_E \| x - \varphi_0(x) \|^4 \; \| g(\varphi_0(x)) \|^2 \; P_\theta(dx)$$

is finite. As in the proof of the proposition 2.3.1, by working conditionally on the radius, we can write with the same notations

$$B = \int_{\mathbb{R}_+} B(R) \rho(dR)$$

with

$$B(R) = \int_E \| x - \varphi_0(x) \|^4 \, \| g(\varphi_0(x)) \|^2 \; U_{R,\theta}(dx) \ .$$

Now it is clear that, for every $R \in \mathbb{R}_+$,

$$B(R) \; \leq \; R^2 \int_{S_{R,\theta}} \| x - \varphi_0(x) \|^2 \, \| g(\varphi_0(x)) \|^2 \; U_{R,\theta}(dx) \ .$$

Hence, if we use the condition (H3) stated in the remark 2.2.1, we obtain

$$B(R) \; \leq \; \alpha R^2 \int_{S_{R,\theta}} \frac{\| x - \varphi_0(x) \|^2}{\| \varphi_0(x) \|^2} \; U_{R,\theta}(dx)$$

$$= \alpha R^2 \, \frac{n-k}{k} \int_{S_{R,\theta}} \frac{\| x - \varphi_0(x) \|^2 / n-k}{\| \varphi_0(x) \|^2 / k} \; U_{R,\theta}(dx)$$

As $U_{R,\theta}$ is spherically symmetric around $\theta$, it is well known (cf. [13] and [14]) that, under that distribution, the statistic $(\| X - \varphi_0 \|^2 / n-k)/(\| \varphi_0 \|^2/k)$ has a distribution independent of R which is a singly noncentral distribution of Fisher with $n-k$ and k degrees of freedom (cf. [9]). Then the last integral is the first moment of this distribution which is finite for $k \geq 3$.

Therefore, for some positive constant L,

$$B \; \leq \; L \int_{\mathbb{R}_+} R^2 \; \rho(dR) \ .$$

Using the same decomposition of the integrations with respect to a spherically symmetric distribution, the finiteness of the risk of $\varphi_0$ assures that the last integral in that inequality is finite (this is in fact condition (H1) according to the remark 2.2.1).

## 5. Appendix

The following lemma is a critical step in our argument.

**Lemma**

Let (E, < , >) be an n-dimensional euclidian space and let $\Theta$ be a k-dimensional linear subspace of E. Let $\varphi_0$ be the orthogonal projector from E onto $\Theta$.

For a positive number R and a vector $\theta$ in $\Theta$, let $U_{R,\theta}$ be the uniform distribution on the sphere $S_{R,\theta} = \{x \in E / \| x \| = R\}$ (where $\| \;\|$ is the norm connected with the inner product < , >).

Then the image by $\varphi_0$ of $U_{R,\theta}$ is absolutely continuous with respect to the Lebesgue measure on $\Theta$ and a density is given by

$$t \to C_R^{n,k} \left(R^2 - \| t - \theta \|^2\right)^{\frac{n-k}{2} - 1}$$

where

$$C_R^{n,k} = \frac{\Gamma\left(\frac{n}{2}\right) R^{n-1}}{\Gamma\left(\frac{n-k}{2}\right) \pi^{k/2}}$$

is the normalization coefficient of that density.

Proof

This result is stated by Kelker [10]. A proof can be found in Cellier and Fourdrinier [7].

### Acknowledgments

The authors would like to thank Professor M.T. Wells for his stimulating discussions and valuable comments which lead to improving the earlier version of this paper. It was partly written when the second author was visiting Cornell University.

## References

[1] Berger, J.O. (1975). Minimax estimation of location vectors for a wide class of densities. *Ann. Statist.* 3, No. 6, 1318-1328.

[2] Berger, J.O. (1976). Admissible minimax estimation of a multivariate mean with arbitrary quadratic loss. *Ann. Statist.* 4, No. 1, 223-226.

[3] Bock, M.E. (1975). Minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* 3, No. 2, 209-218.

[4] Brandwein, A.C. and Strawderman, W.E. (1991). Generalizations of James-Stein estimators under spherical symmetry. *Ann. Statist.*, Vol 19, No 3, 1639-1650.

[5] Brandwein, A.C. and Strawderman, W.E. (1991). Improved estimates of location in the presence of an unknown scale. *J. Multivariate Anal.* 39, 305-314.

[6] Brown, L.D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* 37, No. 5, 1087-1136.

[7] Cellier, D. and Fourdrinier, D. (1990). Sur les lois à symétrie elliptique. Séminaire de Probabilités XXIV, 1988/89. Lecture Notes in Mathematics. Springer-Verlag.

[8] Cellier, D., Fourdrinier, D. and Robert, C. (1989). Robust shrinkage estimators of the location parameter for eliptically symmetric distributions. *J. Multivariate Anal.* 29, 39-52.

[9] Johnson, N.L. and Kotz, S. (1970). Distributions in Statistics: Continuous Univariate - Distributions 2. Wiley, New York.

[10] Kelker, D. (1970). Distribution theory of spherical distribution and a location-scale parameter generalization. Sankhyā Ser.A32, 419-430.

[11] Kruskall, W. (1961). The coordinate-free approach to Gauss-Markov and its application to missing and extra observations. In "Proceedings, Fourth Berkeley Symp. Math. Statist. Probab., 1", pp. 635-451. Univ. of California Press, Berkeley.

[12] Kruskall, W. (1968). When are Gauss-Markov and least squares estimators the same? A coordinate-free approach. *Ann. Math. Statist.* 39, 70-75.

[13] Muirhead, R.J. (1982). Aspects of multivariate statistical theory. Wiley, New York.

[14] Philoche, J.L. (1977). Une condition de validité pour le test F. Statistique et Analyse des Données 1, 37-60.

[15] Scheffé, H. (1959). The analysis of variance. Wiley, New York.

[16] Stein, C. (1956). Inadmissibility of the usual estimator of the mean of a multivariate normal distribution. In "Proceedings, Third Berkeley Symp. Math. Statist. Probab., 1," pp. 197-206. Univ. of California Press, Berkeley.

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

[18] Stone, M. (1987). Coordinate-free Multivariate Statistics. Oxford Univ. Press (Clarendon), London/New York.

Analyse et Modèles Stochastiques
URA CNRS 1378
Mathématiques UFR des Sciences
Université de Rouen
B.P. 118
76134 MONT SAINT AIGAN Cedex