Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 8, 2013, pages 561–575.
DOI: http://dx.doi.org/10.4314/afst.v8i1.5

Quadratic loss estimation of a location parameter when a subset of its components is unknown

Idir Ouassou[†,1] **and Mustapha Rachdi**[‡]

[†]Université Cadi Ayyad, Ecole Nationale des Sciences Appliquées, Av. Abdelkrim Khattabi, BP. 575, Marrakech, Maroc
[‡]Univ. Grenoble-Alpes, Laboratoire AGIM FRE 3405 CNRS, Equipe GEM, Université de P. Mendès-France de Grenoble, UFR SHS, BP. 47, 38040 Grenoble Cedex 09, France

Received: October 01, 2013; Accepted: November 15, 2013

Copyright © 2013, Journal Afrika Statistika. All rights reserved

Abstract. We consider the problem of estimating the quadratic loss $||\delta-\theta||^2$ of an estimator δ of the location parameter $\theta = (\theta_1, \ldots, \theta_p)$ when a subset of the components of θ are restricted to be nonnegative. First, we assume that the random observation X is a Gaussian vector and, secondly, we suppose that the random observation has the form (X, U) and has a spherically symmetric distribution around a vector of the form $(\theta, 0)$ with $\dim X = \dim\theta = p$ and $\dim U = \dim 0 = k$. For these two settings, we consider two location estimators, the least square estimator and a shrinkage estimators, and we compare theirs unbiased loss estimators with improved loss estimator.

Résumé. On considère le problème de l'estimation du coût quadratique $||\delta - \theta||^2$ d'un estimateur δ de la moyenne $\theta = (\theta_1, \ldots, \theta_p)$ d'une loi à symétrie sphérique lorsque l'on sait que certaines composantes θ_i de celle-ci sont positives ou nulles. En premier lieu, lorsque X est un vecteur gaussien, on s'intéresse à l'amélioration de l'estimateur sans biais λ^0 de $||\delta - \theta||^2$ par des estimateurs de la forme $\lambda^0(X) + h(X)$ en fournissant des conditions sur la fonction h. On étend ensuite cette problématique à un modèle distributionnel où l'on dispose d'un vecteur résiduel U : la loi de (X, U) est supposée à symétrie sphérique autour du couple $(\theta, 0)$ et l'on considère des estimateurs de coût de la forme $\lambda^0(X) + ||U||^4 h(X)$.

Key words: Spherical symmetry; Quadratic loss; Least square estimator; Unbiased loss estimator; James-Stein estimation; Minimaxity.
AMS 2010 Mathematics Subject Classification : Primary 62C20; 62C15; 62C10.

Idir Ouassou : idir.ouassou@ensa.ac.ma
Mustapha Rachdi: Mustapha.Rachdi@upmf-grenoble.fr, mustapha.rachdi@agim.eu
[1]This work was supported financially by the Académie des Sciences Hassan II from Morocco

1. Introduction

In this paper we are interested by the estimation of the loss incurred when using the least square estimator and an improved estimator of the location parameter of a spherically symmetric distribution, when a subset of the components of this parameter are restricted to some constraints. The assumption on our model is that the random observation has the form (X, U) and has a spherically symmetric distribution around a vector of the form $(\theta, 0)$ with $\dim X = \dim \theta = p$ and $\dim U = \dim 0 = k$. In that model, the p-dimensional part of the location, say θ, is unknown and the k-dimensional part of the observation is the residual vector. First we construct, for unbiased estimator of the loss, a dominating shrinkage-type estimator, in terms of squared-error loss. These results complement these of James-Stein-type estimation of a location parameter. An important feature of our results is that the proposed loss estimators dominate the unbiased estimator for the entire class of spherically symmetric distributions under constraints. That is, the domination results are robust with respect to the spherical symmetry distribution under constraints.

Notice that the problem of estimating the loss (without any constraint on the scale parameter) was first considered by Lehmann and Scheffé (1993), who estimated the power of a statistical test. Recently, Johnstone (1988), Lele (1993) and Fourdrinierand Wells (1995) have discussed this problem in a variety of situations.

In this paper, we remain in the contexte of spherically symmetric distribution where a residual vector U is available. The type of constraints we consider are that all or only a subset of the θ_i are restricted to be nonnegative. Recall tat Fourdrinierand Wells (1995), Ouassou and Rachdi (2011) have studied this problem but without any constraint on θ and they give an estimators class of the form $\lambda^0 - c||U||^4/||X||^2$ which dominate the unbiased quadratic loss estimator $\lambda^0(X) = p||U||^2/k$ of the minimax estimator X.

Suppose we wish to estimate θ, by a decision rule $\delta(X, U)$ using the sum of squared-error loss $||\delta(X, U) - \theta||^2$. This loss is unobservable since it depends on θ; hence one may wish to estimate it by $\lambda(X, U)$ from the data. To study how well λ estimates the loss, a further distance measure is needed; for mathematical convenience, we use squared error to evaluate $\lambda(X)$. Thus the risk incurred by this latter λ is:

$$R(\lambda, \delta, \theta) = E_\theta \left[\left(\lambda(X, U) - ||\delta(X, U) - \theta||^2 \right)^2 \right].$$

We say that a loss estimator λ_1 dominates λ_2 if $R(\lambda_1, \theta, \delta) < R(\lambda_2, \theta, \delta)$.

In section 2, we study the quadratic loss estimator of an estimator of the scale parameter $\theta = (\theta_1, \ldots, \theta_p)$ when some of its components θ_i are nonnegative. First, when the random vector X is normally distributed with mean θ and of covariance matrix the identity I_p. In this setting, the random vectors X and U are independent then the situation becomes the same as when just the vector X is available, so here, we do not consider the residual vector U.

So in the first step, we study the quadratic loss estimation $||\delta_0(X) - \theta||^2$ of the maximum likelihood estimator $\delta_0(X) = (\delta_{01}(X), \ldots, \delta_{0p}(X))$ where, for all i ($1 \leq i \leq p$),

$$\delta_{0i}(X) = \begin{cases} \max(0, X_i) & \text{for } i = 1, \ldots, s \\ X_i & \text{for } i = s+1, \ldots, p \end{cases}$$

We get the unbiased estimator $\lambda^0(X)$ of its quadratic loss, which we then we improve by giving some domination conditions for the estimators of the form $\lambda^0(X) + h(X)$.

Afterward, we consider a more general class of estimators of θ of the form $\delta_g(X) = \delta_0(X) + g(X)$. In this case, we give an unbiased estimator $\lambda_g^0(X)$ of the quadratic loss $||\delta_g(X) - \theta||^2$ of the estimator $\delta_g(X)$ and we establish the conditions for which $\lambda_g^0(X)$ is dominated by an estimator of the form $\lambda_g^h(X) = \lambda_g^0(X) + h(X)$.

Finally, still for the spherically symmetric distribution context where a residual vector U is available, we estimate the quadratic loss $||\delta_g(X, U) - \theta||^2$ of any estimator of the form $\delta_g(X, U) = \delta_0(X) + ||UU||^2 g(X)$ of the scale parameter θ, by giving an unbiased estimator $\lambda_g^0(X)$ of this quadratic loss and conditions for which $\lambda_g^0(X)$ is dominated by a more general estimator of the form $\lambda_g^h(X) = \lambda_g^0(X) + ||U||^4 h(X)$.

2. The Gaussian model framework

Assume that a p-dimensional random vector $X = (X_1, \ldots, X_p)$ is observed which is normally distributed with mean vector $\theta = (\theta_1, \ldots, \theta_p)$ such that it has s nonnegative components and covariance identity matrix I_p. Without loss of generality, we assume that $\theta_1 \geq 0, \ldots, \theta_s \geq 0$ for $s \leq p$.

2.1. Quadratic loss estimation of the maximum likelihood estimator

In this case, the maximum likelihood estimator is $\delta_0(X) = (\delta_{01}(X), \ldots, \delta_{0p}(X))$ where

$$\delta_{0i}(X) = \begin{cases} \begin{cases} 0 & \text{if } X_i < 0 \\ X_i & \text{if } X_i \geq 0 \end{cases} & \text{for } i = 1, \ldots, s \\ X_i & \text{for } i = s+1, \ldots, p \end{cases}$$

Il is convenientl to write it on the form: $\delta_0(X) = X + \gamma(X)$ where $\gamma(X) = (\gamma_1(X), \ldots, \gamma_p(X))$ with

$$\gamma_i(X) = \begin{cases} \begin{cases} -X_i & \text{if } X_i < 0 \\ 0 & \text{if } X_i \geq 0 \end{cases} & \text{for } i = 1, \ldots, s \\ 0 & \text{for } i = s+1, \ldots, p. \end{cases}$$

An unbiased estimator of the quadratic loss $||\delta_0(X) - \theta||^2$ of the usual estimator $\delta_0(X)$ is given by:

$$\lambda^0(X) = p + \sum_{i=1}^{s} (X_i^2 - 2) 1_{[X_i < 0]}. \tag{1}$$

where 1_A denotes the indicator function on the set A. In fact, the risk of $\delta_0(X)$ in θ may be written as follows:

$$\begin{aligned} E_\theta\left[||\delta_0(X) - \theta||^2\right] &= E_\theta\left[||X - \theta + \gamma(X)||^2\right] \\ &= E_\theta\left[||X - \theta||^2\right] + E_\theta\left[||\gamma(X)||^2\right] + 2E_\theta\left[(X - \theta)^t \gamma(X)\right] \end{aligned}$$

This decomposition is valid since $E_\theta\left[||X-\theta||^2\right] = p < +\infty$ (which is the risk of X) and that:

$$E_\theta\left[||\gamma(X)||^2\right] = E\left[\sum_{i=1}^{s} X_i^2 1_{[X_i<0]}\right] \leq E_\theta\left[||X||^2\right] < +\infty$$

Then, the Schwartz inequality provides that $E_\theta\left[(X-\theta)^t\gamma(X)\right]$ exist.
As the function γ is weakly differentiable, then the Lemma 5 allows to voice this expectation:

$$E_\theta\left[(X-\theta)^t\gamma(X)\right] = E_\theta\left[\text{div}\ \gamma(X)\right] = -E\left[\sum_{i=1}^{s} 1_{[X_i<0]}\right]$$

by some simple calculation of the divergence of $\gamma(X)$.
Finally the risk of $\delta_0(X)$ is finite and may be written such that:

$$p + E_\theta\left[\sum_{i=1}^{s}(X_i^2-2)1_{[X_i<0]}\right].$$

An alternative class of estimators which improve $\lambda^0(X)$ are of the form:

$$\lambda^h(X) = \lambda^0(X) - h(X).$$

where h is some function defined from $\mathbb{R}^p$ into $\mathbb{R}$. The difference in risk between $\lambda^h(X)$ and $\lambda^0(X)$ is given by Lemma 1 below. In Lemma 1, as in all what follows, the weakly differentiability of the function h is assumed because it is a natural hypothesis as in Stein Lemma (cf. Stein; 1981 and Johnstone; 1988). This assumption permits also to give some examples for which the differentiability is not valid.

Lemma 1. *For every weakly differentiable function $h : \mathbb{R} \to \mathbb{R}^p$ and for every $\theta \in \mathbb{R}^p$, the difference in risks of $\lambda^h(X)$ and $\lambda^0(X)$ is equal to:*

$$\begin{aligned} &R\left(\lambda^h,\delta_0,\theta\right) - R\left(\lambda^0,\delta_0,\theta\right) \qquad (2)\\ &= E_\theta\left[h^2(X) + 4sh(X) + 2\triangle h(X)\right] - 4E_\theta\left[h(X)\sum_{i=1}^{s}\left(1_{[X_i>0]} + (X_i^2 - X_i\theta_i)1_{[X_i<0]}\right)\right]\end{aligned}$$

provided these expectation exist, where $\triangle$ (resp. div) denotes the Laplacian operator (resp. divergency).

Proof. For a fixed $\theta \in \mathbb{R}^p$, we have that:

$$\begin{aligned} &R\left(\lambda^h,\delta_0,\theta\right) - R\left(\lambda^0,\delta_0,\theta\right)\\ &= E_\theta\left[(\lambda^0(X) - h(X) - ||\delta_0(X)-\theta||^2)^2 - (\lambda^0(X) - ||\delta_0(X)-\theta||^2)^2\right]\\ &= E_\theta\left[h^2(X) - 2(\lambda^0(X) - ||\delta_0(X)-\theta||^2)h(X)\right].\end{aligned}$$

Since,

$$\begin{aligned} &\lambda_0(X) - ||\delta_0(X)-\theta||^2\\ &\qquad = p - s + \sum_{i=1}^{s} I_{[X_i\geq 0]} + \sum_{i=1}^{s}(X_i^2-1)1_{[X_i<0]} - \sum_{i=1}^{s}(\delta_{0i}(X)-\theta_i)^2 - \sum_{i=s+1}^{p}(\delta_{0i}(X)-\theta_i)^2\\ &\qquad = (p-2s) + 2\sum_{i=1}^{s}\left(1_{[X_i>0]} + (X_i^2 - X\theta_i)I_{[X_i<0]}\right) - ||X-\theta||^2.\end{aligned}$$

then,

$$E_\theta\left[(\lambda^0(X)-||\delta_0(X)-\theta||^2)h(X)\right] \tag{3}$$
$$= (p-2s)E_\theta\left[h(X)\right]+2\sum_{i=1}^{s}E_\theta\left[\left(1_{[X_i>0]}+(X_i^2-X_i\theta_i)1_{[X_i<0]}\right)h(X)\right]$$
$$-E_\theta\left[||X-\theta||^2h(X)\right].$$

and, by using the Lemma 6 result on the function h, we get

$$E_\theta\left[||X-\theta||^2h(X)\right]=pE_\theta\left[h(X)\right]+E_\theta\left[\triangle h(X)\right]. \tag{4}$$

Combining the results above, (3) and (4) , the difference in risk equals for every θ

$$E_\theta\left[h^2(X)+4sh(X)+2\triangle h(X)\right]-4E_\theta\left[h(X)\sum_{i=1}^{s}\left(1_{[X_i>0]}+(X_i^2-X_i\theta_i)1_{[X_i<0]}\right)\right].$$

□

As a first application of the above we have the following result.

Theorem 1. *Let h be a two-times weakly differentiable and positive function on $\mathbb{R}^p$. If the function h satisfy*

$$h^2+2\triangle h+4sh<0 \tag{5}$$

then, the estimator $\lambda^h(X)$ dominates $\lambda^0(X)$, for $\theta\in\mathbb{R}_+^s\times\mathbb{R}^{p-s}$.

Proof. The Condition (5) is a consequence of the that the second expectation of the equation (2) is negative. Indeed, we have by hypothesis that, for all $i=1,\dots,s$, $\theta_i>0$ and therefore $-X_i\theta_i1_{[X_i<0]}\geq 0$ for $i=1,\dots,s$. Then we have

$$\sum_{i=1}^{s}(1_{[X_i>0]}+(X_i^2-X_i\theta_i)1_{[X_i<0]})\geq 0$$

and so

$$E_\theta\left[h(X)\sum_{i=1}^{s}(1_{[X_i>0]}+(X_i^2-X_i\theta_i)1_{[X_i<0]})\right]\geq 0$$

since the function h is assumed positive, which is the desired result. □

Remark 1. The case which there is no positivity constraint on the components of the parameter $\theta\mathbb{R}^p$ may be result in the fact that $s=0$. Note that in this interpretation, the inequality (5) coincides with the condition given by Johnstone (1988).

2.2. Quadratic loss estimation of a general estimator

We consider the estimaton of the loss of a class of shrinkage estimators δ, estimators incurred when using the maximum likelihood estimator $\delta_0(X)$ as

$$\delta(X) = \delta_g(X) = \delta_0(X) + g(X)$$

where g is weakly differentiable function from $\mathbb{R}^p$ to $\mathbb{R}^p$.

As before we first develop an unbiased estimator of the loss function $||\delta_g(X) - \theta||^2$ of the estimator $\delta_g(X)$. Indeed, by application of Lemma 5 to the function $\gamma + g$, we can write

$$\begin{aligned} E_\theta\left[||\delta_g(X) - \theta||^2\right] &= E_\theta\left[||X + \gamma(X) + g(X) - \theta||^2\right] \\ &= E_\theta\left[||X - \theta + \gamma(X) + g(X)||^2\right] \\ &= E_\theta\left[||X - \theta||^2 + 2(X-\theta)(\gamma(X) + g(X)) + ||\gamma(X) + g(X)||^2\right] \\ &= p + E_\theta\left[||\gamma(X) + g(X)||^2 + 2\text{div }(\gamma + g)(X)\right], \end{aligned}$$

therefore an unbiased estimator of $||\delta_g(X) - \theta||^2$ has the following form

$$\lambda_g^0(X) = p + ||\gamma + g||^2 + 2\text{div }(\gamma + g).$$

In the signal of this section we focus to prove the domination of the unbiased estimator λ_0 of the quadratic loss by computing estimators λ_g^h of the form

$$\lambda_g^h(X) = \lambda_g^0(X) - h(X)$$

for some real function unknown h. Then in Lemma 2 we develop the difference in risk between $\lambda_g^h(X)$ and $\lambda_g^0(X)$.

Lemma 2. *For every twice-weakly differentiable function $h : \mathbb{R} \to \mathbb{R}^p$ and a weakly differentiable function $g : \mathbb{R}^p \to \mathbb{R}^p$. For every $\theta \in \mathbb{R}^p$ the difference in risk between $\lambda_g^h(X)$ and $\lambda_g^0(X)$ is given by:*

$$R\left(\lambda_g^h, \delta_g, \theta\right) - R\left(\lambda_g^0, \delta_g, \theta\right) = E_\theta\left[h^2(X) + 2 \triangle h(X) + 4\left\langle \triangledown h(X), \gamma(X) + g(X)\right\rangle\right].$$

Proof. Let $\theta \in \mathbb{R}^p$. The difference in risk between $\lambda_g^h(X)$ and $\lambda_g^0(X)$ is equal to

$$\begin{aligned} &R\left(\lambda_g^h, \delta_g, \theta\right) - R\left(\lambda_g^0, \delta_g, \theta\right) \\ &\quad = E_\theta\left[(\lambda_g(X) - ||\delta_g(X) - \theta||^2)^2 - (\lambda_g^0(X) - ||\delta_g(X) - \theta||^2)^2\right] \\ &\quad = E_\theta\left[(\lambda_g^0(X) - h(X) - ||\delta_g(X) - \theta||^2)^2 - (\lambda_g^0(X) - ||\delta_g(X) - \theta||^2)^2\right] \\ &\quad = E_\theta\left[h^2(X)\right] - 2E_\theta\left[(\lambda_g^0(X) - ||\delta_g(X) - \theta||^2)h(X)\right]. \end{aligned}$$

It is clear, by using Lemma 6 to the function h, that the second expectation in this last expression can be expressed as follows:

$$\begin{aligned} &E_\theta\left[(\lambda_g^0(X) - ||\delta_g(X) - \theta||^2)h(X)\right] \\ &\quad = E_\theta\left[\left(p + ||\gamma(X) + g(X)||^2 + 2\text{div }(\gamma + g)(X) - ||X + \gamma(X) + g(X) - \theta||^2\right)h(X)\right] \\ &\quad = E_\theta\left[\left(p + 2\text{div }(\gamma + g)(X) - 2(X - \theta)^t(\gamma + g)(X) - ||X - \theta||^2\right)h(X)\right] \\ &\quad = 2E_\theta\left[h(X)\text{div}g(X) - \text{div}(g \cdot h)(X) - \frac{1}{2} \triangle h(X)\right] \\ &\qquad + 2E_\theta\left[(\text{div}\gamma - (X - \theta)^t\gamma(X))h(X)\right]. \end{aligned}$$

Otherwise as

$$\begin{aligned}\text{div}\gamma(X) - (X-\theta)^t\gamma(X) &= -\sum_{i=1}^{s} I_{[X_i<0]} + \sum_{i=1}^{s}(X_i-\theta_i)X_i 1_{[X_i<0]} \\ &= -s + \sum_{i=1}^{s} 1_{[X_i\geq 0]} + (X_i^2 - X_i\theta_i)1_{[X_i<0]}.\end{aligned}$$

Then we are now able to give an expression for the difference in risk

$$\begin{aligned}&E_\theta \Big[h^2(X) + 2\triangle h(X) + 4\text{div}(g\cdot h)(X) - 4h(X)\cdot \text{div}g(X) + 4sh(X) \\ &\quad -4h(X)\sum_{i=1}^{s} 1_{[X_i>0]} + (X_i^2 - X_i\theta_i)1_{[X_i<0]}\Big] \\ &= E_\theta\left[h^2(X) + 2\triangle h(X) + 4\langle \bigtriangledown h(X), \gamma(X)+g(X)\rangle\right],\end{aligned}$$

which achieves the proof of this Lemma. □

The following theorem gives a condition of domination of the estimator $\lambda_g^0(X)$ by $\lambda_g^h(X)$.

Theorem 2. *Let h be a positive function twice-weakly differentiable on $\mathbb{R}^p$ and $p\geq 4$. A sufficient condition under which the estimator $\lambda_g^h(X)$ dominates the unbiased estimator $\lambda_g^0(X)$, for all $\theta\in R_+^s\times R^{p-s}$, is that h satisfies de differential inequality*

$$h^2 + 2\triangle h + 4sh + 4\langle\bigtriangledown h, g\rangle \leq 0.$$

Proof. Demonstration follows the same lines as that of Theorem 1. □

3. The spherically unimodal framework

We consider the model introduced in section1 where (X,U) is a $p+k$ random vector having a symmetrically distribution around de $p+k$ vector $(\theta,0)$ Here $\dim X = \dim\theta = p$ and $\dim U = \dim 0 = k$.

For the constraints on θ, we takes those in paragraph 2 where some components of the mean θ is assumed positive (i.e. $\theta_i\geq 0$ for $i=1,\ldots,s$ with $s\leq p$).

We will study the estimation of quadratic loss $||\delta_g(X,U)-\theta||^2$ for any estimator of the form $\delta_g(X,U) = X+\gamma(X)+||U||^2 g(X)$ of the mean θ, where g is a function twice-weakly differentiable on $\mathbb{R}^p$ in $\mathbb{R}^p$. General conditions for such an estimator $\delta_g(X,U)$ dominates the estimator $X+\gamma(X)$ are given in Fourdrinier et al. (2003), Brandwein and Strawderman (1991), Brandwein *et al.* (1991) and Cellier *et al.* (1989).

The next lemma give an unbiased estimator of the loss $||\delta_g(X,U)-\theta||^2$ of the estimator $\delta_g(X,U)$.

Lemma 3. *For every weakly differentiable function $g:\mathbb{R}^p\to\mathbb{R}^p$ and for all $\theta\in\mathbb{R}^p$, an unbiased estimator of the loss function $||\delta_g(X,U)-\theta||^2$ of the estimator $\delta_g(X,U)$ is equal to*

$$\lambda_g^0(X,U) = p\frac{||U||^2}{k} + 2\frac{||U||^2}{k}\text{div}\gamma(X) + \frac{2}{k+2}||U||^4\text{div}g(X) + ||\gamma(X)+||U||^2 g(X)||^2. \quad (6)$$

Proof. let $\theta \in \mathbb{R}^p$ fixed. They can write

$$\begin{aligned}
&E_\theta\left[||\delta_g(X,U)-\theta||^2\right] = E_\theta\left[||X+\gamma(X)+||U||^2g(X)-\theta||^2\right]\\
&= E_\theta\left[||(X-\theta+\gamma(X))+||U||^2g(X)||^2\right]\\
&= E_\theta\left[||X-\theta+\gamma(X)||^2+2(X-\theta+||U||^2\gamma(X)).g(X)+||U||^4||g(X)||^2\right].
\end{aligned}$$

Frome the Lemma 7, for $\alpha=0$ and $\alpha=1$ respectively, we have

$$E_\theta\left[(X-\theta)^t\gamma(X)\right] = E_\theta\left[\frac{||U||^2}{k}\text{div }\gamma(X)\right]$$

and

$$E_\theta\left[||U||^2(X-\theta)^tg(X)\right] = E_\theta\left[\frac{||U||^4}{k+2}\text{div }g(X)\right].$$

It comes then

$$\begin{aligned}
&E_\theta\left[||\delta_g(X,U)-\theta||^2\right]\\
&= E_\theta\left[p\frac{||U||^2}{k}+2\frac{||U||^2}{k}\text{div }\gamma(X)+||\gamma(X)||^2+2||U||^2\gamma(X)\cdot g(X)\right.\\
&\quad\left.+2||U||^2(X-\theta)^tg(X)+||U||^4||g(X)||^2\right]\\
&= E_\theta\left[p\frac{||U||^2}{k}+2\frac{||U||^2}{k}\text{div }\gamma(X)+||\gamma(X)||^2+2||U||^2\gamma(X)g(X)+||U||^2||g(X)||^2\right.\\
&\quad\left.+\frac{2}{k+2}||U||^4\text{div }g(X)\right]\\
&= E_\theta\left[p\frac{||U||^2}{k}+2\frac{||U||^2}{k}\text{div }\gamma(X)+\frac{2}{k+2}||U||^4\text{div }g(X)+||\gamma(X)+||U||^2g(X)||^2\right]
\end{aligned}$$

then the result. □

Remark 2. Where there is no positivity constraint on the components of θ can be interpreted by taking $\gamma\equiv 0$. We thus find the unbiased estimator given by Cellier and Fourdrinier (1995).

In the following of this section, we show that the estimator $\lambda_g^0(X,U)$ can be improved through a more general class of estimators of the form

$$\lambda_g^h(X,U) = \lambda_g^0(X,U)+||U||^4h(||X||^2) \tag{7}$$

where $h(\cdot)$ is a negative real-valued function and twice weakly differentiable.

Order to compare two estimators $\lambda_G^h(X,U)$ and $\lambda_g^0(X,U)$ we needs to calculate the difference in their risk

$$\Delta R = R\left(\lambda_g^h(X,U),\delta_g(X,U),\theta\right)-R\left(\lambda_g^0(X,U),\delta_g(X,U),\theta\right).$$

It is the object of the following lemma.

Lemma 4. *For every weakly differentiable function* $g : \mathbb{R}^p \to \mathbb{R}^p$*, for every twice weakly differentiable function* $h : \mathbb{R}^+ \to \mathbb{R}$ *and for all* $\theta \in \mathbb{R}^p$*, an unbiased estimator of the difference in risk between* $\lambda_g^h(X,U)$ *and* $\lambda_g^0(X,U)$ *is equal to*

$$||U||^8 \left(h^2(||X||^2) - \frac{2}{(k+4)(k+6)} \triangle h(||X||^2) \right.$$
$$\left. + \frac{4}{k+2} h(||X||^2) \text{div } g(X) - \frac{4}{k+6} \text{div } \left(h(||X||^2) g(X) \right) \right)$$
$$+ \frac{4}{k+4} ||U||^6 \left(\frac{2p}{k} h(||X||^2) - \langle \triangledown h(||X||^2), \gamma(X) \rangle + \frac{4}{k} h(||X||^2) \text{div } \gamma(X) \right). \quad (8)$$

Proof. See Appendix.

As a first application of the previous result, we have the following theorem where we assume that the distribution of (X,U) is unimodal is to say that it have a density of the form $(x,u) \to G(||x-\theta||^2 + ||u||^2)$ where $G(\cdot)$ is a nonincreasing function.

Theorem 3. *We assume that the distribution of* (X,U) *is unimodal. If the function* h *satisfies the following conditions*

1. *is a negative function, twice weakly differentiable and concave.*
2. $\frac{\partial h}{\partial X_i}(||X||^2) \le -\frac{h(||X||^2)}{||X||^2} X_i$ *for all* $i = 1, \ldots, p$ *then the estimator* $\lambda_g^h(X,U)$ *dominates the unbiased estimator* $\lambda_g^0(X,U)$*, for* $\theta \in \mathbb{R}^s \times \mathbb{R}^{p-s}$*, since provided that*

$$h^2(||X||^2) - \frac{2}{(k+4)(k+6)} \triangle h(||X||^2) - \frac{4}{k+6} \text{div } (g \cdot h)(X) + \frac{4}{k+2} h(||X||^2) \text{div } g(X)$$
$$- \frac{2s}{(k+4)(k-2)} \frac{h(||X||^2)}{||X||^2} \le 0 \quad (9)$$

provided these expectation exist.

Proof. Using Lemma 4, we get that the difference in risk

$$E_\theta \left[||U||^8 \left(h^2(||X||^2) - \frac{2}{(k+4)(k+6)} \triangle h(||X||^2) \right. \right.$$
$$\left. \left. - \frac{4}{k+6} \text{div } (g \cdot h)(X) + \frac{4}{k+2} h(||X||^2) \text{div } g(X) \right) \right] + \frac{4}{k+4} \left(\frac{2p}{k} E_\theta \left[||U||^6 h(||X||^2) \right] \right.$$
$$\left. + \frac{4}{k} E_\theta \left[||U||^6 h(||X||^2) \text{div } \gamma(X) \right] - E_\theta \left[||U||^6 \langle \triangledown h(||X||^2), \gamma(X) \rangle \right] \right). \quad (10)$$

Using the results of the paper Fourdrinier et al. (2003) (demonstration of theorem 3.1), we show that the last expectation of (10) is bounded by

$$- \frac{s}{k-2} E_\theta \left[||U||^8 \frac{h(||X||^2)}{||X||^2} \right]. \quad (11)$$

Indeed,

$$\begin{aligned}-E_\theta\left[||U||^6\left\langle\bigtriangledown h(||X||^2),\gamma(X)\right\rangle\right] &= E_\theta\left[||U||^6\sum_{i=1}^s\frac{\partial h(||X||^2)}{\partial X_i}X_iI_{[X_i<0]}\right]\\ &\le -2\sum_{i=1}^s E_\theta\left[||U||^6\frac{h(||X||^2)}{||X||^2}X_i^2I_{[X_i<0]}\right]\\ &\le -\frac{s}{k-2}E_\theta\left[||U||^8\frac{h(||X||^2)}{||X||^2}\right]\end{aligned}$$

because the function h satisfies the same conditions as those of function r of theorem 3.1 du Fourdrinier et al. (2003). Then

$$E_\theta\left[-\frac{4}{k+4}||U||^6\left\langle\bigtriangledown h(||X||^2),\gamma(X)\right\rangle\right]\le\frac{-2s}{(k+4)(k-2)}E_\theta\left[||U||^8\frac{h(||X||^2)}{||X||^2}\right]. \tag{12}$$

To conclude we need Lemma 9 to show that, the sum of the second and the third expectation of the expression (10) is negative. Using Lemma 9, we get that,

$$E_\theta\left[||U||^6h(||X||^2)\text{div }\gamma(X)\right]\le-\frac{s}{2}E_\theta\left[h(||X||^2)||U||^6\right]. \tag{13}$$

In effect

$$\begin{aligned}E_\theta\left[||U||^6h(||X||^2)\text{div }\gamma(X)\right] &= -E_\theta\left[||U||^6h(||X||^2)\sum_{i=1}^s I_{[X_i<0]}\right]\\ &= -\sum_{i=1}^s E_\theta\left[||U||^6h(||X||^2)I_{[X_i<0]}\right]\\ &\le -\frac{s}{2}E_\theta\left[h(||X||^2)||U||^6\right]\end{aligned}$$

the last inequality being acquired by using the fact that the function $-h(\cdot)$ is nonnegative and the Lemma 9. So, from equation (13), we have

$$\begin{aligned}&E_\theta\left[\frac{8p}{k(k+4)}||U||^6h(||X||^2)+\frac{16}{k(k+4)}||U||^6h(||X||^2)\,\gamma(X)\right] \qquad (14)\\ &\le E_\theta\left[\frac{8p||U||^6h(||X||^2)}{k(k+4)}-\frac{8s}{k(k+4)}||U||^6h(||X||^2)\right]\\ &= 8(p-s)E_\theta\left[\frac{||U||^6}{k(k+4)}h(||X||^2)\right]\le 0\end{aligned}$$

since the function $h(\cdot)$ is negative and $p-s\ge 0$.

Finally we acquire as upper bound of the difference in risk

$$E_\theta\left[||U||^8\left(h^2(||X||^2)-\frac{2}{(k+4)(k+6)}\bigtriangleup h(||X||^2)-\frac{4}{k+6}(g\cdot h)(X)+\frac{4}{k+2)}h(||X||^2)\text{div }g(X)-\frac{2s}{k(k+4)(k-2)}\frac{h(||X||^2)}{||X||^2}\right)\right],$$

what gives the desired result. □

Remark 3. Where there is no positivity constraint on the components of θ can be interpreted by taking $s = 0$ and $\gamma \equiv 0$. We thus find the condition of domination of Theorem 3.1 in Fourdrinier et al. (2003).

The following corollary gives an example of a function h and function g which satisfies the conditions of Theorem 3.

Corollary 1. *Assume that $p > 4$, $p > \frac{s+8}{2}$, $k > 2 + \frac{4s}{2p-8-s}$ and that the distribution is unimodal. For the functions $g(X) = -\frac{d}{||X||^2}X$ and $h(||X||^2) = -\frac{c}{||X||^2}$ a sufficient condition under which the estimator $\lambda_g^h(X,U)$ dominates the unbiased estimators $\lambda_g^0(X,U)$, for all $\theta \in \mathbb{R}^s \times \mathbb{R}^{p-s}$, is that*

$$0 < c < \frac{4(p-4)}{k+2} + 4\left(\frac{(p-4)}{k+6} - \frac{(p-2)}{k+2}\right)d - \frac{2s}{(k+4)(k-2)}. \tag{15}$$

where the constante d satisfy

$$0 \leq d \leq \frac{1}{2}\frac{(k+2)(p-\frac{s+8}{2})\left(k-\left[2+\frac{4s}{2p-s-8}\right]\right)}{(k+4)(k-2)(2p+k-2)} \tag{16}$$

Proof. Let us consider the usual shrinkage estimator g of θ with shrinkage factor g defined by $g(X) = d\frac{X}{||X||^2}$, where d is a positive constant, and the shrinkage loss estimator used in the previous corollary with the shrinkage function h defined by $h(t) = -c\frac{X}{||X||^2}$ where c is a positive constante, for every $X \in \mathbb{R}^p$, it easy to check that

$$\triangle h(||X||^2) = \sum_{i=1}^{p}\frac{\partial^2 h}{\partial^2 X_i}(||X||^2) = 2c\frac{p-4}{||X||^4},$$

$$\text{div } g(X) = \sum_{i=1}^{p}\frac{\partial g_i}{\partial^2 X_i}(X) = -d\frac{p-2}{||X||^2}.$$

and

$$\begin{aligned}\text{div } (h\cdot g)(X) &= h(||X||^2)\text{div } g(X) + \left\langle \triangledown h(||X||^2), g(X)\right\rangle \\ &= \frac{c}{||X||^2}\frac{d(p-2)}{||X||^2} + \sum_{i=1}^{p}\frac{\partial h}{\partial X_i}(X)g_i(X) \\ &= cd\frac{p-4}{||X||^4}.\end{aligned}$$

The member of left of the condition of domination (9) becomes

$$\frac{c^2}{||X||^4} - \frac{2}{(k+4)(k+6)}\frac{2c(p-4)}{||X||^4} - \frac{4}{(k+6)}\frac{cd(p-4)}{||X||^4} + \frac{4}{k+2}\frac{cd(p-2)}{||X||^4} + \frac{2sc}{(k+4)(k-2)}\frac{1}{||X||^4}$$
$$= \frac{c}{||X||^4}\left[c - \frac{4(p-4)}{(k+4)(k+6)} - \frac{4(p-4)}{k+6}d + \frac{4(p-2)}{k+2}d + \frac{2s}{(k+4)(k-2)}\right].$$

Hence the sufficient condition of domination is

$$c - \frac{4(p-4)}{(k+4)(k+6)} - \left(\frac{4(p-4)}{k+6} - \frac{4(p-2)}{k+2}\right) d + \frac{2s}{(k+4)(k-2)} < 0$$

what is identical to

$$0 < c < 4\frac{(p-4)}{(k+4)(k+6)} + 4d\left(\frac{(p-4)}{k+6} - \frac{(p-2)}{k+2}\right) - \frac{2s}{(k+4)(k-2)}.$$

A simple calcultion shows that the right-and side is positive if and only $0 \leq d \leq \frac{1}{2}\frac{(k+2)(p-\frac{s+8}{2})\left(k-\left[2+\frac{4s}{2p-s-8}\right]\right)}{(k+4)(k-2)(2p+k-2)}$ and $k > 2 + \frac{4s}{2p-8-s}$. This completes the proof. □

Remark 4. The optimal of c is given by

$$c^* = \frac{2(p-4)}{(k+4)(k+6)} + 2d\left(\frac{(p-4)}{k+6} - \frac{(p-2)}{k+2}\right) - \frac{s}{(k+4)(k-2)}.$$

It is clear that the optimal shrinkage factor of the loss estimator depends on the shrinkage factor of te point estimate. We found that wih the optimal d, that is, $d = \frac{p-2}{k+2} - \frac{s}{k-2}$, the optimal c is still positive for reasonable values of p and k.

Comment

It is worthwhile to note that the stated condition for domination are too restrictive is certain case ($s \leq 4$). In particular, what is needed is the finiteness of the risk, which is governed by the final term of the expression. Thus, if $s = 0$, it suffices that $p \geq 5$, $k \geq 1$ and $0 < c < 4\frac{(p-4)}{(k+4)(k+6)} + 4d\left(\frac{(p-4)}{k+6} - \frac{(p-2)}{k+2}\right)$ for same $0 \leq d \leq \frac{1}{2}\frac{(p-4)(k+2)}{(2p+k-2)(k+4)}$ which is same bound as in the unrestricted case. If $s = 1$, or $s = 2$, it suffices that $p \geq 5$ and if $s = 4$, that $p \geq 7$. The condition on k for $s \geq 1$ are as in the corollary. In particular, if $s = 1$ and $p = 5$, we require $k \geq 10$ Similarly, if $s = 2$ and $p = 6$, we also require $k \geq 10$ in general for $p = s + 4$ we require $k \geq 10$, but, for $s = 3$ and $p = 6$, we need $k \geq 26$. For $s = 4$ and $p = 7$, we need that $k \geq 8$. For $p = s$ we need to $p \geq 9$. In general, for fixed s, the value of k which is required decreases to 3 as p increases to $\left[\frac{5s+4}{2}\right] + 1$.

4. Appendix

The first two lemma are given respectively by Stein (1981) and by Stein (1981) and Johnstone (1988).

Lemma 5. *For every weakly differentiable function $g : \mathbb{R}^p \to \mathbb{R}^p$, for every $\theta \in \mathbb{R}^p$, we have*

$$E_\theta\left[(X-\theta).g(X)\right] = E_\theta\left[\operatorname{div} g(X)\right] \tag{17}$$

provided these expectations exist.

Lemma 6. *For every weakly differentiable function $g : \mathbb{R}^p \to \mathbb{R}$,for every $\theta \in \mathbb{R}^p$, we have*

$$E_\theta\left[||X-\theta||^2 g(X)\right] = E_\theta\left[\Delta g(X) + p g(X)\right]$$

provided these expectations exist.

Proof of Lemma 6

We using (17) to the function $(X-\theta)g(X)$.

The straightforward proof of the two next lemma is given respectively by Fourdrinier and Strawderman (1996) and by Fourdrinierand Wells (1995).

Lemma 7. *For every weakly differentiable function $g : \mathbb{R}^p \to \mathbb{R}^p$, for every $\theta \in \mathbb{R}^p$ and for every integer α,*

$$E_\theta\left[||U||^{2\alpha}(X-\theta).g(X)\right] = \frac{1}{k+2\alpha}E_\theta\left[||U||^{2\alpha+2}\mathrm{div}\, g(X)\right]$$

Lemma 8. *For every twice weakly differentiable function $g : \mathbb{R}^p \to \mathbb{R}^+$, for every $\theta \in \mathbb{R}^p$ and for every integer q, we have*

$$E_{R,\theta}\left[||U||^q||X-\theta||^2 g(X)\right]$$
$$= \frac{p}{k+q}E_{R,\theta}\left[||U||^{q+2}g(X)\right] + \frac{1}{(k+q)(k+q+2)}E_{R,\theta}\left[||U||^{q+4}\Delta g(X)\right].$$

Lemma 9. *Assume X is a real-valued random variable with symmetric unimodal distribution about $\theta \in \mathbb{R}^+$. If f is a nonnegative function on $\mathbb{R}^+$, then*

$$E_\theta\left[f(X^2)I_{[X<0]}\right] \le \frac{1}{2}E_\theta\left[f(X^2)\right]. \tag{18}$$

□

Proof of Lemma 9

Note that , by the symmetrical unimodal assumption, X has density of the forme $(x,u) \to g\left((x-\theta)^2\right)$ with g nonincreasing function. The lemma reduces showing that

$$E_\theta\left[f(X^2)\left(I_{[X<0]} - \frac{1}{2}\right)\right] \le 0. \tag{19}$$

The left-hand side of (19) can be written as

$$\frac{1}{2}E_\theta\left[f(X^2)\left(I_{[X<0]} - I_{[X\ge 0]}\right)\right]. \tag{20}$$

Conditioning on $|X|$, and noting by E the expectation of the distribution of $|X|$, the expression (20) is

$$\frac{1}{2}E\left[f(X^2)\left(g\left((-|X|-\theta)^2\right) - g\left((|X|-\theta))^2\right)\right)\right].$$

So, by hypothesis we have $\theta > 0$ then

$$(|X|-\theta)^2 \le (-|X|-\theta)^2.$$

By unimodatily of X (the function $g(\cdot)$ is decreasing), it is clear that

$$g\left((-|X|-\theta)^2\right) \le g\left((|X|-\theta))^2\right).$$

and hence that an upper bound of (19) is given by

$$\frac{1}{2}E\left[f(X^2)\left(g\left((-|X|-\theta)^2\right) - g\left((|X|-\theta))^2\right)\right)\right] \le 0$$

what gives the desired result. □

Proof of Lemma 4

For $\theta \in \mathbb{R}^p$ fixed. We can compute using (6) and (7)

$$\Delta R(\theta) = E_\theta\left[\left(\lambda_g^h(X,U) - ||\delta_g(X,U)-\theta||^2\right)^2\right] - E_\theta\left[\left(\lambda_g^0(X,U) - ||\delta_g(X,U)-\theta||^2\right)^2\right]$$

$$= E_\theta\left[\left(\lambda_g^0(X,U) + ||U||^4 h(||X||^2) - ||\delta_g(X,U)-\theta||^2\right)^2 - \left(\lambda_g^0(X,U) - ||\delta_g(X,U)-\theta||^2\right)^2\right]$$

$$= E_\theta\left[||U||^8 h^2(||X||^2) + 2p\frac{||U||^6}{k}h(||X||^2) - 2||U||^4||X-\theta||^2 h(||X||^2)\right]$$

$$+4E_\theta\left[\frac{||U||^6}{k}\left(\text{div}\gamma(X) + \frac{k}{k+2}||U||^2\text{div}g(X)\right)h(||X||^2)\right]$$

$$-4E_\theta\left[||U||^4(X-\theta)\left(\gamma(X)+||U||^2 g(X)\right)h(||X||^2)\right].$$

Using Lemma 8, for $q=4$ and $g(X)=h(||X||^2)$, we obtain

$$E_\theta\left[||U||^4||X-\theta||^2 h(||X||^2)\right]$$

$$= \frac{p}{k+4}E_\theta\left[||U||^6 h(||X||^2)\right] + \frac{1}{(k+4)(k+6)}E_\theta\left[||U||^8 \triangle h(||X||^2)\right]$$

and using Lemma 7, for the function $(\gamma + ||U||^2 g)h$, we get

$$E_\theta\left[||U||^4(X-\theta)\gamma(X)h(||X||^2)\right] = \frac{1}{k+4}E_\theta\left[||U||^6\text{div}\left(\gamma(X)h(X)\right)\right]$$

and

$$E_\theta\left[||U||^6(X-\theta)g(X)h(||X||^2)\right] = \frac{1}{k+6}E_\theta\left[||U||^8\text{div}\left(g(X)h(X)\right)\right]$$

Then the difference in risk be comme

$$E_\theta\left[||U||^8 h^2(||X||^2) + 2p\frac{||U||^6}{k}h(||X||^2) + 4\frac{||U||^6}{k}h(||X||^2)\left(\gamma(X) + \frac{k}{k+2}||U||^2 g(X)\right)\right.$$

$$-\frac{2}{(k+4)(k+6)}||U||^8 \triangle h(||X||^2) - \frac{4}{k+4}||U||^6\text{div}\left(\gamma(X)h(X)\right)$$

$$\left.-2\frac{p}{k+4}||U||^6 h(||X||^2) - \frac{4}{k+6}||U||^8\text{div}\left(g(X)h(X)\right)\right]$$

$$= E_\theta\left[||U||^8\left\{h^2(||X||^2) - \frac{2}{(k+4)(k+6)}\triangle h(||X||^2)\right.\right.$$

$$\left.\left.+\ \frac{4}{k+2}h(X)\text{div }\left(g(X)\right) - \frac{4}{k+6}\text{div }\left(h(X)g(X)\right)\right\}\right]$$

$$+E_\theta\left[||U||^6\left\{\frac{2p}{k}h(||X||^2) - \langle\triangledown h(||X||^2), \gamma(X)\rangle + \frac{4}{k}h(||X||^2)\text{div }\gamma(X)\right\}\right].$$

So the unbiased estimator desired is

$$||U||^8\left(h^2(||X||^2) - \frac{2}{(k+4)(k+6)}\triangle h(||X||^2) + \frac{4}{k+2}h(X)\left(g(X)\right) - \frac{4}{k+6}\left(h(X)g(X)\right)\right)$$

$$+4||U||^6\left(2\frac{p}{k}h(||X||^2) - \frac{1}{k+4}\text{div}(h(X)\gamma(X)) + \frac{1}{k}h(X)\text{div}\gamma(X)\right)$$

then the desired result. □

References

Brandwein, A. C. and Strawderman, W.E., 1991. Generalizations of James-Stein estimators under spherical symmetry. *The Annals of Statistics.* **19**(3): 1639-1650.

Brandwein, A. C., Ralescu, S. and Strawderman, W.E., 1991. Shrinkage estimators of the location parameter for certain spherically symmetric distributions. *Ann. Inst. Statist. Math.* **45**(3), 551-563.

Cellier, D. and Fourdrinier, D., 1995. Shrinkage estimators under spherical symmetry for the general linear model. *Journal of Multivariate Analysis.* **52**(2): 338-351.

Cellier, D., Fourdrinier, D. and Robert, C., 1989. Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *Journal of Multivariate Analysis.* **29**: 39-52.

Fourdrinier, D. and Wells, M.T., 1995. Estimation of a loss function for spherically symmetric distribution in the general linear model. *The Annals of Statistics.* **23**(2): 571-592.

Fourdrinier, D. and Strawderman, W.E., 1996. A paradox concerning shrinkage estimators : should a known scale parameter be replaced by an estimated value in the shrinkage factor? *Journal of Multivariate Analysis.* **59**(2), 109-140.

Fourdrinier, D., Ouassou, I. and Strawderman, W., 2003. Estimation of a components vector when some parameters are restricted. *Journal of Multivariate Analysis.* **86**, 14-24,

Johnstone, I., 1988. On inadmissibility of some unbiased estimates of loss. *In Statistical Decision Theory and Related Topics IV (S. S. Gupta and J.O.Berger, Eds.), Academic Press, New York.* **1**, 361–379.

Lehmann, E. L. and Scheffé, H., 1950. Completeness, similar regions and unbiased estimates. *Sankhyā.* **10**: 305-340,

Lele, C., 1993. Admissibility results in loss estimation. *The Annals of Statistics.* **21**, 378-390.

Ouassou, I. and Rachdi, M., 2009. Stein type estimation of the regression operator for functional data. *Advances and Applications in Statistical Sciences.* **1**, no 2, 233-250.

Stein, C., 1981. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics.* **9**, 1135-1151.

Ouassou, I. and Rachdi, M. 2011. Estimation du coût quadratique quand certaines composantes du paramètre de position sont positives C. R. Acad. Sci. Paris, Sér. 1, Volume 349, Issue 17-18, Pages 995-998.