Statistics and Probability Letters 96 (2015) 241–246

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On improved shrinkage estimators for concave loss

Tatsuya Kubokawa [a], Éric Marchand [b], William E. Strawderman [c,*]

[a] Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[b] Université de Sherbrooke, Département de Mathématiques, Sherbrooke Qc, Canada, J1K 2R1
[c] Rutgers University, Department of Statistics and Biostatistics, 501 Hill Center, Busch Campus, NJ 08855, Piscataway, USA

ARTICLE INFO

Article history:
Received 2 July 2014
Received in revised form 23 September 2014
Accepted 24 September 2014
Available online 2 October 2014

MSC:
primary 62C20
62F15
secondary 62A15

Keywords:
Location parameters
Minimaxity
Stein estimation
Concave loss

ABSTRACT

We consider minimax shrinkage estimation of location for spherically symmetric distributions under a concave function of the usual squared error loss. Scale mixtures of normal distributions and losses with completely monotone derivatives are featured.

1. Introduction

This paper concerns minimax shrinkage estimation of a location vector of a spherically symmetric distribution under a loss function which is a concave function of the usual squared error loss. The main contribution is an improvement in shrinkage constants for minimax estimators over those in Brandwein and Strawderman (1980, 1991), and Brandwein et al. (1993), particularly for variance mixtures of normals (and somewhat more generally), and for concave functions of squared error loss for which the derivative of the concave function is completely monotone (and somewhat more generally). For Baranchik-type estimators and for scale mixtures of multivariate normal distributions, we also show that our minimax improvements hold in dimension 3 which improves over the restriction that $p \geq 4$ in the earlier papers.

Specifically, let X have the p-dimensional spherically symmetric density

$$f(\|x-\theta\|^2), \quad x, \theta \in \mathbb{R}^p, \tag{1.1}$$

and consider the problem of estimating the unknown location vector θ under the loss function

$$L(\theta, d) = l(\|d-\theta\|^2), \tag{1.2}$$

where $l(t)$ is a non-negative, non-decreasing concave function, and $\|d-\theta\|^2$ is the usual squared error loss function.

* Corresponding author.
E-mail addresses: tatsuya@e.u-tokyo.ac.jp (T. Kubokawa), eric.marchand@usherbrooke.ca (É. Marchand), straw@stat.rutgers.edu (W.E. Strawderman).

http://dx.doi.org/10.1016/j.spl.2014.09.024
0167-7152/

For a multivariate normal distribution with squared error loss, James and Stein (1961), Baranchik (1970), Strawderman (1971), Stein (1981) and others gave shrinkage estimators of the mean vector which are minimax and which improve on the usual estimator, X, when the dimension, p, is at least three. Strawderman (1974) gave extensions to variance mixtures of normals. Extensions to wider classes of spherically symmetric distributions were provided by Berger (1975), Brandwein and Strawderman (1978), Brandwein (1979), Brandwein and Strawderman (1991) and others.

Brandwein and Strawderman (1980), Brandwein and Strawderman (1991) and Brandwein et al. (1993) gave minimax shrinkage estimators which improve on X in higher dimensions for concave functions of squared error loss.

A basic tool in much of the literature on concave loss is the following simple result which will be used extensively in the following.

Lemma 1.1. *Suppose that X is distributed as in (1.1), and that loss is given by (1.2), where $l(t)$ is a non-negative, non-decreasing concave function such that $l'(t)$ exists.*

a. *Then the risk, $R(\theta,\delta)$, of an estimator of the form $\delta(X)=X+g(X)$, satisfies the inequality*

$$R(\theta,\delta)\le R(\theta,X)+E_\theta[l'(\|X-\theta\|^2)(\|g(X)\|^2-2(X-\theta)'g(X))],\tag{1.3}$$

where the expectation E_θ is with respect to the density (1.1).

b. *Hence $\delta(X)$ dominates X under loss (1.2) if it dominates X under quadratic loss, $\|d-\theta\|^2$, for a location family with density $f^*(\|x-\theta\|^2)$ proportional to $f(\|x-\theta\|^2)l'(\|x-\theta\|^2)$.*

Proof. Part a follows easily, on taking expectations, from the concave function inequality $l(t+y)\le l(t)+yl'(t)$ with $t=\|X-\theta\|^2$ and $y=\|g(X)\|^2+2(X-\theta)'g(X)$.

Part b follows immediately from part a. □

Remark 1.1. That the usual estimator X is minimax follows fairly easily when the underlying spherically symmetric density (1.1) is unimodal and when the loss function (1.2) is monotone (but not necessarily concave) in $\|d-\theta\|^2$. This follows because X is the unique minimum risk equivariant (MRE) estimator under these assumptions and is hence minimax by the well known result that if a minimax estimator exists in a location problem, there is an equivariant minimax estimator. For completeness we formalize this in Theorem 1.1.

Theorem 1.1. *Suppose that the density $f(\|X-\theta\|^2)$ is unimodal and that $\ell'(t)\ge 0$ with $\ell'(t)>0$ on an interval. Then, the usual estimator X is the unique minimum risk equivariant (MRE) estimator, and is hence minimax.*

Proof. Note first that a general equivariant estimator is of the form $X+d$ for some vector d in $\Re^p$, which because of the spherical symmetry we may take to be of the form $(a,0,\dots,0)$. Let $Y=(X_2,\dots,X_p)$, let $\ell_Y((X_1-a)^2)=\ell((X_1-a)^2+\|Y\|^2)$. Then, conditioning on Y, it suffices to show that $E[\ell_Y((X_1-a)^2)|Y]>E[\ell_Y(X_1^2)|Y]$ for $a\ne 0$, where E represents the expectation with respect to the conditional distribution of X_1 given Y. Let $f_Y((x_1-a)^2)$ be the conditional density of X_1 given Y. The conditional risk difference is expressed as

$$\begin{aligned}\Delta_Y(a)&=E[\ell_Y((X_1-a)^2)|Y]-E[\ell_Y(X_1^2)|Y]=E\Big[\int_0^a\frac{d}{dt}\ell_Y((X_1-t)^2)dt|Y\Big]\\&=2\int_{-\infty}^{\infty}\int_0^a(t-x)\ell_Y'((x-t)^2)f_Y(x^2)dtdx\\&=2\int_{-\infty}^{\infty}\int_0^a(-z)\ell_Y'(z^2)f_Y((z+t)^2)dtdz\\&=2\int_{-\infty}^{\infty}(-z)\ell_Y'(z^2)\int_z^{a+z}f_Y(u^2)dudz,\end{aligned}$$

where we used the transformations $z=x-t$ $(dz=dx)$ and $u=t+z$ $(du=dt)$. Let $F_Y(z)=\int_{-\infty}^z f_Y(u^2)du$. Then, we can rewrite the conditional risk difference as

$$\begin{aligned}\Delta_Y(a)&=2\int_{-\infty}^{\infty}z\ell_Y'(z^2)\{F_Y(z)-F_Y(z+a)\}dz\\&=2\int_0^{\infty}z\ell_Y'(z^2)\Big[\{F_Y(z)-F_Y(z+a)\}-\{F_Y(-z)-F_Y(-z+a)\}\Big]dz\\&=2\int_0^{\infty}z\ell_Y'(z^2)\Big[\{F_Y(z)-F_Y(-z)\}-\{F_Y(z+a)-F_Y(-z+a)\}\Big]dz\\&=2\int_0^{\infty}z\ell_Y'(z^2)\Big[\int_{-z}^{z}f_Y(u^2)du-\int_{-z+a}^{z+a}f_Y(u^2)du\Big]dz.\end{aligned}$$