On High Breakdown Point Estimators of Location and Scale in the Multidimensional Case*

Maya Z. Marincheva and Dimitar L. Vandev†

Abstract

In their short communication Neykov and Neychev (1990) have proposed a robustified version of maximum likelihood principle – RML. It leads to two families of robust estimators LME - the Least Median of log density values Estimator, and LTE - the Least Trimmed log - likelihood Estimator. This paper studies the possibility to extend the results of Vandev and Neykov (1993) to a more general (than the multidimensional normal) elliptical family of density functions.

1 Introduction

Neykov and Neychev (1990) proposed a robustified version of maximum likelihood principle – RML. Some of the widely known robust estimators of multivariate location and scatter matrix follow easily from this principle. Among them are MVE - the Maximum Volume Ellipsoide, and MCE - the Minimum Covariance Determinant introdused by Rousseeuw (1986). It is shown (Lopuhaa and Rousseeuw, 1991) that in the Gaussian case they both have a breakdown point of 1/2 - the best that can be achieved.

In this paper we focus our atterntion to a general elliptical family defined by fixed "shape" function $\varphi(z)$. Vandev (1992) developed a technique for computing the breakdown point of LME and LTE. He proved that the breakdown point is not less than $(n-k)/n$, where k is a tuning constant of the estimators which can be chosen by the user within some reasonable range of values. Vandev and Neykov (1993) based on these results studied the connection of the finite – sample breakdown point, dimensionality of the Gaussian distribution and the notion of d – fullness introduced by Vandev (1992).

Our considerations as an extention of the normal case, follow the similar tecnique when proving the statements. We obtain a high breakdown point for LME and LTE when $\varphi(z)$ has a "propriate behavior".

[1]This paper is partly financed by EC Project COST 228
and by The National Fondation of Scientific Investigation, grant MM - 440/94

[2]Postal address: Computer Stochastics, Institute of Mathematics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria.

The main result is that when $\varphi(z)$ is a positive, decreasing, restricted from above function and when $\varphi(z) = O(e^{-\alpha z})$, the set of density values for the sample form a $(p+1)$ – full set of function of the unknown scale parameter the matrix S. This implies that RML - estimators $LME(k)$ and $LTE(k)$ have breakdown point not less than $(n-k)/n$, for k - expected number of outliers being the parameter of the estimator.

2 Definitions and Notations

Let consider $x_1, x_2, \ldots, x_n$- a sample of n independent observations in the p-dimentional euclidean space E^p, over a random value ξ with the following density function:

$$f(x, \mu, S) = \frac{Cp}{\sqrt{det}(S)} \varphi((x-\mu)' S^{-1}(x-\mu)).$$

Here Cp is a standardized constant, and μ and S denote the location and scale parameters correspondingly.

Our aim is to find high breakdown point robust estimators for the unknown parameters. Vandev (1993) showed that the breakdown point of both $LME(k)$ and $LTE(k)$ estimators is not less than $(n-k)/n$ if the set of n positive functions $\{-\ln f(x_i, \mu, S), i \in \{1, 2, \ldots, n\}\}$ is d – full and $(n+d)/2 \le k \le (n-d)$.

In order to apply this result, it only remains to determine the conditions that the function $\varphi(x)$ must satisfy, as well as the value of d for the density family mentioned above.

First of all we should recall the basic definition introdused by Neykov and Neychev (1990) and later extended a little bit by Vandev and Neykov (1993) concerning RML.

Definition 1: The Least k-ordered of log density Estimator (LME) of θ for $k > \frac{n}{2}$ is defined as:

$$LME(k)(x_1, x_2, \ldots, x_n) = arg \min_{\theta}(-\ln f(x_{l(k)}, \mu, S)),$$

The Least Trimmed log - likelyhood Estimator (LTE) of θ is defined as:

$$LTE(k)(x_1, x_2, \ldots, x_n) = arg \min_{\theta} \sum_{i=1}^{k} (-\ln(f(x_{l(i)}, \mu, S))),$$

where $f(x_{l(1)}, \mu, S) \ge f(x_{l(2)}, \mu, S) \ge \ldots \ge f(x_{l(n)}, \mu, S)$ are the ordered density values and θ denote the unknown parameter.

Definition 2: The real valued function $g(z)$ defined on a topological space Z is called subcompact, if its Lebesque sets $L(M) = \{z : g(z) \le M\}$ are compact or empty for all constants M.

Definition 3: A finite set F of n functions is called d – full, if for each subset of cardinality d of F, the supremum of all functions in this subset is a subcompact function.

3 Basic results

Let $F = \{-\ln f(x_1, \mu, S), -\ln f(x_2, \mu, S), \ldots, -\ln f(x_n, \mu, S)\}$.

Firstly we should determine the conditions which $\varphi(x)$ must satisfy. In order to apply Vandev's results we should obtain the positivity of the functions $-\ln f(x_i, \mu, S)$. For this it is necessary that $\varphi(x)$ be a positive, decreasing, restricted from above function.

The only restriction on parameters is that there must exist $\epsilon > 0$ such that $det(S) \geq \epsilon$.

Theorem 1 *If $x_1, x_2, \ldots, x_n$ is a sample with a density function*

$$f(x, \mu, S) = \frac{Cp}{\sqrt{det(S)}} \varphi((x - \mu)' S^{-1} (x - \mu)),$$

then the finite set F form

(1) a (p+1) – full set with probability 1, if the scale parameter S is unknown, and $\varphi(x)$ satisfies the extra assumption $\varphi(z) = O(e^{-\alpha z}), \alpha \geq 0$;

(2) a 1 – full set if only the location parameter μ is unknown.

For the proof we needed the following lemmas.

Lemma 1 *(a)For any $(p + 1)$ points in general position $I(\mu, S) = \max_{i \in \{1,2,\ldots,p+1\}}(-\ln f(x_i, \mu, S))$ is subcompact in μ and S if the scale parameter S is unknown and $\varphi(z) = O(e^{-\alpha z})$;*

(b) $I(\mu, S) = -\ln f(x_i, \mu, S)$ is subcompact in μ, if only the location μ is unknown.

Lemma 2 *If $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of BS^{-1} where*

$$B = \frac{1}{p + 1} \sum_{i=1}^{p+1} (x_i - \overline{x})(x_i - \overline{x})',$$

then:

$$e^{-H} \leq \lambda_i \leq \frac{eH}{e - 1},$$

where $H = \sum_{i=1}^{p} \lambda_i - \ln \prod_{i=1}^{p} \lambda_i$

Lemma 3 *(a standard fact from Linear Algebra) For α_i – the eigenvalues of S, if real constants exist α and β, such that from $\alpha \leq \alpha_i \leq \beta$ follows: $\alpha \leq \|S\| \leq \beta$*

Proof of the theorem:

Let us suppose that $\varphi(z)$ is an uninterrupted on the left lunction.

Case (2): We consider the case when only μ is an unknown parameter. Then for an arbitrary real constant K let denote as

$$\begin{aligned} I(\mu) &= \{\mu : -lnf(x_i, \mu, S) \leq K\} \\ &= \left\{\mu : \varphi((x_i - \mu)' S^{-1} (x_i - \mu)) \geq C\right\}, \end{aligned}$$

where $C := e^{K1} = const > 0, \quad K1 := -K - \ln \frac{C_p}{\sqrt{detS}} = const.$

We must show that $I(\mu)$ is a compact function in μ.

(I) Let there exist a point z_0 such that : $\varphi(z_0) = C$. Then there exists an internal J, such that $\varphi(z) = C$.

(II) Let suppose that no point z_0, exists for which $\varphi(z_0) = C$,and denote as J the following interval:

$$J := \{z : \varphi(z) > C\}$$

In these both cases because $\varphi(z)$ is positive and decreasing it turns out that J is an interval restricted on the right.

Therefore $\forall \ \ z \in J : \varphi(z) \geq \varphi(sup(J))$.

We must pay attention to the fact that the last statement is satisfied because $\varphi(z)$ is uninrerrupted on the left.

Now it is not hard to extend $I(\mu)$ to the set $I1(\mu)$, definited as:

$$I_1(\mu) = \left\{\mu : (x_i - \mu)' S^{-1} (x_i - \mu) \leq sup(J)\right\}.$$

As $I_1(\mu)$ is a restricted set we can make the conclusion that $I(\mu)$ is restricted as well.

In case (1) we introduce:

$$\begin{aligned} I(\mu, S) &= \max_{i \in \{1,2,\ldots,p+1\}} \{-\ln f(x_i, \mu, S)\} \\ &= -\ln Cp + \frac{1}{2} \ln(detS) - \ln \varphi(\max_{i \in \{1,2,\ldots,p+1\}} ((x_i - \mu)' S^{-1} (x_i - \mu))). \end{aligned}$$

and denote by A :

$$\begin{aligned} A &:= \{(\mu, S) : I(\mu, S) \leq K\} \\ &= \{(\mu, S) : \frac{1}{2} \ln(detS) - \ln \varphi \max_{i \in \{1,2,\ldots,p+1\}} ((x_i - \mu)' S^{-1} (x_i - \mu)) \leq K1\}. \end{aligned}$$

where $K1 = K + \ln Cp$.

Using the inequalities:

$$\max_{i\in\{1,2,\ldots,p+1\}}((x_i-\mu)'S^{-1}(x_i-\mu)) \geq \frac{1}{p+1}\sum_{i=1}^{p+1}(x_i-\mu)'S^{-1}(x_i-\mu)$$

and

$$\frac{1}{p+1}\sum_{i=1}^{p+1}((x_i-\mu)'S^{-1}(x_i-\mu)) \geq \frac{1}{p+1}\sum_{i=1}^{p+1}((x_i-\overline{x})'S^{-1}(x_i-\overline{x})),$$

where $\overline{x}$ is the mean of $x_1, x_2, \ldots, x_{p+1}$, the set A expands to the set C, which is:

$$C := \left\{(\mu, S) : \frac{1}{2}\ln(detS) - \ln\varphi\left(\frac{1}{p+1}\sum_{i=1}^{p+1}(x_i-\overline{x})'S^{-1}(x_i-\overline{x})\right) \leq K1\right\}.$$

Denoting by B and Z correspondingly $B := \frac{1}{p+1}\sum_{i=1}^{p+1}(x_i-\overline{x})(x_i-\overline{x})'$ and $Z := S^{-1}$, we can finally come to $C := \{(\mu, S) : \sqrt{(}det(BZ))\varphi(Tr(BZ)) \geq L\}$, with $L = e^{-K1}.\sqrt{(}detB) = const.$

Let $\lambda_1{}', \lambda_2{}', \ldots, \lambda_p{}'$ and $\alpha_1, \alpha_2, \ldots \alpha_p$ are the eigenvalues of BZ and S corespondengly.

Let denote as: $\lambda_i = \frac{\lambda_i{}'}{coeff}$ for $i \in \{1, 2, \ldots, p\}$, where $coeff$ is arbitrary positive.

Then C turns into:

$$C = \left\{(\mu, S) : \frac{1}{2}\left(\sum_{i=1}^{p}\lambda_i - \ln\prod_{i=1}^{p}\lambda_i\right) \leq \ln\varphi(coeff^p.\sum_{i=1}^{p}\lambda_i + \frac{1}{2}.\sum_{i=1}^{p}\lambda_i{}' - \ln L\right\}.$$

For $H := \sum_{i=1}^{p}\lambda_i - \ln\prod_{i=1}^{p}\lambda_i$, we can expand C to C_1 :

$$C_1 := \left\{S : H \leq 2.\left(\sum_{i=1}^{p}\lambda_i + \ln\varphi\left(coeff^p.\sum_{i=1}^{p}\lambda_i\right) - \ln L\right)\right\}$$

Using the extra assumption $\varphi(z) = O(e^{-\alpha z})$, finally we manage one more time to extend C_1 to $C_2 := \{S : H \leq M - \sum_{i=1}^{p}\lambda_i(1 - coeff^p.\alpha)\}$.

Because by appropriate $coeff$ we can make $(1 - coeff^p.\alpha) > 0$, and obtain:

$$C_2 \subset D := \{S : H \leq M\}.$$

From Lemma2 we obtain that $e^{-M} \leq \lambda_i \leq \frac{eM}{e-1}$, for $i \in \{1, 2, \ldots, p\}$ and multiplying all these inequalities reminding that $det(S) = \prod_{i=1}^{p}\alpha_i$ we obtain the follwing double inequality:

$$\frac{(e-1)^p.b}{e^pM^p} \leq \prod_{i=1}^{p}\alpha_i \leq e^{pM}.b,$$

where $detB = b = const$.

Therefore there are positive constants α and β, such that $\alpha \leq \alpha_i \leq \beta \quad \forall i \in \{1, 2, \ldots, p\}$, and proceeding in the same way we obtain $F := \{S : \alpha \leq \alpha_i \leq \beta, \quad i \in \{1, 2, \ldots, p\}\}$ and $A \subset C \subset C_1 \subset D \subset F$.

From $\alpha \leq \alpha_i \leq \beta$ and Lemma3 we obtain that $\alpha \leq \|S\| \leq \beta$, which is equivalent to the fact that F and therefore A is restricted too.

Now it only remains to study that A is a closed set.

Let us consider $\{(\mu_n, S_n)\}, (\mu_n, S_n) \in A \; \forall n \in N$, where $\mu_n \rightarrow \mu$ and $S_n \rightarrow S$ when $n \rightarrow \infty$.

We shall show that $(\mu, S) \in A$.

For this it is convenient to denote as

$$Z_n := \max_{i \in \{1,2,\ldots,p+1\}} ((x_i - \mu_n)' {S_n}^{-1} (x_i - \mu_n))$$

and as

$$Z := \max_{i \in \{1,2,\ldots,p+1\}} ((x_i - \mu)' S^{-1} (x_i - \mu)),$$

where $\mu_n = \mu$, in case of μ – known.

Let us assume that $(\mu, S) \notin A$, e.g. $\varphi(Z) < C$ and let us consider the row $\{Z_n\}$.

In these conditions let us assume as well that there exist $k \in N$, such that $y_k > y$.

After a few rows we obtain contradictions with the both assumptions, and we finally obtain, that A is a closed set. But above we proved that A is restricted, therefore A is a compact set.

Unfortunately it turns out that when the extra assumption: $\varphi(z) = O(e^{-\alpha z})$ is not satisfied e.g. in the particular case when $\varphi(z) = z^{-\alpha z}$, the RML principle is not obtainable.

The gap when $\varphi(z)$ behaved itself between $z^{-\alpha z}$ and $e^{-\alpha z}$ is an open question.

We must pay attention to the fact that from a density viewpoint, the restriction for uninterruptedness on the left for $\varphi(z)$ is purely technical. Terefore this restriction for $\varphi(z)$, need not be satisfied in the sense of equivalent density functions.

References

[1] Neykov, N.M.& Neytchev, P.N. (1990) "A Robust Alternative of the ML estimators"
COMPSTAT'90, Short communications, Dubrovnik, Yugoslavia, pp. 99 – 100.

[2] Vandev, D.L.(1993). "A Note on the Breakdown Point of the Least Median and Least Trimmed Estimators", *Statistics and Probability Letters,* 16, pp. 117 – 119.

[3] Vandev, D.L.& Neykov, N.M. (1993) "Robust Maximum Likelihood in the Gaussian Case", *New Directions in Statistical Data Analysis and Robustness*, pp. 259 – 264.

[4] Hampel, F.R. Ronchetty, E.M. Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics: The Approach Based on influance Functions,* John Wiley and Sons New York.

[5] Rousseeuw, P.J. (1986). "Multivariate Estimation with High Breakdown Point". In: *Mathematical Statistics and Applicaations,* Vol. B, W.Grossman, G.Pflg, I.Vincze and W.Wertz (eds.), Dordrecht: Reidel Publishing Company, pp. 283-297.