Asymptotic properties of location estimators based on projection depth

Jeankyung Kim
Inha University, Statistics
253 Yonghyundong, Namgu
Incheon, Korea
jkkim@anova.inha.ac.kr

Jinsoo Hwang
Inha University, Statistics
253 Yonghyundong, Namgu
Incheon, Korea
jshwang@anova.inha.ac.kr

1. Introduction

The recent paper of Bai and He(2000) established asymptotic distributions of maximal depth regression estimator and the half-space median, which is characterized through a max-min operation of a continuous Gaussian process. They also mentioned that Tyler(1994) gave another example with the same type of limiting distribution. In this paper, we give a rigorous proof of the limiting distribution of projection median using the result of Bai and He(2000) under a general setting. We also show $\sqrt{n}$ consistency of the trimmed mean and a metrically trimmed mean based on the projection based distance.

2. Limit distribution of the projcetion Median

Let $X_1, \ldots, X_n$ be a sequence of independent observations from a fixed distribution P on $\Re^d$ and let P_n be the empirical measure which gives mass n^{-1} to each $X_1, \ldots, X_n$. Let Ω be a sample space of $\boldsymbol{X_n} = (X_1, \ldots, X_n)$, that is, $\Omega = (\Re^d)^n$ and $I\!P = P^n$ be the corresponding probability measure. We adopt linear functional notation for integration and set notation for indicator function.

Let S be a $(d-1)$ dimensional subspace of $\Re^d$ such that $S = \{u \in \Re^d : ||u|| = 1\}$ where $||\cdot||$ is the Euclidean norm. Then the sample projection median is defined as the solution to the min-max problem:

$$\theta_n = \arg \inf_{x \in \Re^d} \sup_{u \in S} \frac{|u'x - Med_{1 \le i \le n}(u'X_i)|}{MAD_{1 \le i \le n}(u'X_i)},$$

and the population analogous is $\theta_0 = \arg\inf_{x\in\Re^d}\sup_{u\in S}\frac{|u'x-Med(u'X)|}{MAD(u'X)}$. Without loss of generality, we consider the case of $\theta_0 = 0$.

¿From the fact that $Med_{1\le i\le n}(u'X_i) = O_p(n^{-1/2})$ and $(MAD_{1\le i\le n}(u'X_i) - MAD(u'X)) = O_p(n^{-1/2})$ uniformly in $u \in S$, it is easy to show that $\theta_n = O_p(n^{-1/2})$. Once $\theta_n = O_p(n^{-1/2})$ is established, we need to consider the rescaled estimator,

$$\tau_n \;=\; \sqrt{n}\theta_n \;=\; \arg\inf_{x\in\Re^d}\sup_{u\in S} W_n(x,u),$$

where

$$W_n(x,u) \;=\; \frac{|u'x - \sqrt{n}Med_{1\le i\le n}(u'X_i)|}{MAD_{1\le i\le n}(u'X_i)}, \quad \text{for } x\in\Re^d, u\in S. \tag{1}$$

Let P^u be the one dimensional marginal distribution of $u'X$, and p^u be the corresponding density function. We need some assumptions on the underlying distribution.

(P1) P^u has an unique median at 0 for all unit vector u.

(P2) P^u has a bounded density p^u, with $p^u(0) > 0$, and $p^u(s)$ is continuous in u and at $s = 0$.

(P3) For any $\gamma \in S$, $(u/MAD(u'X))'\gamma$ is maximized if and only if $u = \gamma$.

Lemma 1 *Under (P1)-(P2), $\{W_n\}$ defined in (**??**) converges in distribution to the process*

$$W(x,u) \;=\; \frac{|u'x - G(u)|}{MAD(u'X)},$$

where G is a centered Gaussian process with continuous sample path and covariance kernel

$$H(u,v) = \frac{1}{4p^u(0)p^v(0)} P g_u(\cdot) g_v(\cdot),$$

where $g_u(x) = sign(u'x)$.

Lemma 2 *Under (P1)-(P3), for each sample path of W, solution of* $\arg\inf_{x\in\Re^d}\sup_{u\in S} W(x,u)$ *is unique.*

Theorem 1 *Under (P1)-(P3),*

$$\tau_n \;=\; \sqrt{n}\,\theta_n \to \arg\inf_{x\in\Re^d}\sup_{u\in S} W(x,u).$$

REFERENCES

Bai, Z and He, X (1999), Asymptotic Distribution of the Maximal Depth Estimators for Regression and Multivariate Location, *The Annals of Statistics* **27**, 1616-1637.

Tyler, D.E.(1994), Finite sample breakdown points of projection based multivariate location and scatter statistics, *Annals of Statistics* **22**, 1024-1044.

RESUME

Nous établissons les propriétés asymptotiques des estimateurs de l'emplacement basées sur la profondeur de la projection. La distribution limitée de la projection moyenne est obtenue en utilisant le résultat de Bai et He(2000) en conditions normales. Les taux $\sqrt{n}$ de convergence de la moyenne corrigée et de la moyenne corrigée métriquement sont obtenus.