MIN-MAX BIAS ROBUST REGRESSION

YAN LIU

1. Reference

1.1. **流れ-1.**

Martin, Yohai and Zamar (1989), AS.
Yohai and Zamar (1993), AS.
Berrendero and Zamar (2001), AS.
Berrendero Mendes and Tyler (2007), AS.

1.2. **流れ-2.**

Donoho and Liu (1988a), AS
Donoho and Liu (1988b), AS
He and Simpson (1993), AS

2. Fundamental Setting

- Klotz gave a matrix version of Szegö's

2.1. **Definition.**

(1) g is

$$g(s, \|\theta\|) = E_{H_0}\rho\Big(\frac{y - x^{\mathrm{T}}\theta}{s}\Big)$$

and the inverse of g_1 and g_2 are defined by,

(g_1) the inverse of g with respect to s;

(g_2) the inverse of g with respect to $\|\theta\|$.

note. つまり、$g_1(a, 0)$ は回帰真値の周りでの量で、$g_2(a, b)$ は...

Date: July 23, 2015.

(2)

$$V_\epsilon = \{H; H = (1-\epsilon)H_0 + \epsilon H^*\}$$

(3) The asymptotic bias of T at H

$$b(T, H) = \|T(H)\|$$

and correspondingly,

$$B_\epsilon(T) = \sup\{\|T(H)\|; H \in V_\epsilon\}$$

note. 何らかの推定量で、$B_\epsilon(T)$ を最小にすれば、これが minimax となる。

3. Conditions

3.1. **on ρ.** Let ρ be a real-valued function on $\mathbb{R}$ satisfying the following conditions.

3.2. **when ρ is a jump function.**

$$\rho_c(u) = \begin{cases} 0, & |u| < c, \\ 1, & |u| \geq c. \end{cases}$$

3.3. **S-estimate functional.** Let $s(\theta, H) = s(F^\theta)$, where

$$s(F) = \inf\{s > 0; E_F \rho\left(\frac{u}{s}\right) \leq b\}.$$

note. b と設定した時の、F における最小の influence function のスケールを s としている。

$T(H)$ is said to be an S-estimate functional of regression if there exists a sequence $\theta_n \in \mathbb{R}^p$ such that

$$\lim_{n\to\infty} \theta_n = T(H)$$

and

$$\lim_{n\to\infty} s(\theta_n, H) = \inf_{\theta\in\mathbb{R}^p} s(\theta, H).$$

Lemma 3.1. *If ρ satisfies A1 and H satisfies*

$$\sup_{\|\theta\|=1} P_H(x^{\mathrm{T}}\theta = 0) < 1 - b,$$

then there exist some sequence θ_n and $T(H)$ to be an S-estimate.

Consequently, if ρ is additionally continuous,

$$s(T(H), H) = \min\{s(\theta, H); \theta \in \mathbb{R}^p\}.$$

3.4. **Contrast function.** Set $g(s,\|\theta\|)$ as

$$g(s,\|\theta\|) = E_{H_0}\rho\Big(\frac{y - x^{\mathrm{T}}\theta}{s}\Big)$$

Lemma 3.2. (i) *ρ satisfies A1;*
(ii) *F_0 satisfies A2;*
(iii) *G_0 satisfies A3.*
Then g is continuous, strictly increasing w.r.t. $\|\theta\|$; strictly decreasing in s for $s > 0$.

3.5. **Assumption A.**

(A1) Symmetric and nondecreasing on $[0,\infty)$, with $\rho(0) = 0$.
Bounded, with $\lim_{u\to\infty}\rho(u) = 1$.
ρ has only a finite number of discontinuities.
note. ρ は分布関数もどき。

(A2) F_0 is absolutely continuous with density f_0 which is symmetric, continuous and strictly decreasing for $u < 0$.

(A3) G_0 is spherical and $P_{G_0}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\theta} = 0)$ for all $\theta \in \mathbb{R}^p$ with $\theta \neq 0$.

4. Main Results

Lemma 4.1. *Assume that ρ_1 satisfies A1, F_0 satisfies A2 and G_0 satisfies A3. Then g is continuous, strictly increasing with respect to $\|\boldsymbol{\theta}\|$ and strictly decreasing in s for $s > 0$.*

note. これがいえると、漸近的な話でなくとも成り立つ。

Proof. (i) To show $\forall H = (1-\epsilon)H_0 + \epsilon H^*$,

$$\|\theta\| > c \Rightarrow s(\theta, H) > s(0, H).$$

note. s とは何か。

□

Theorem 4.2. *Under Lemma, $B_\epsilon(T)$ over V_ϵ is*

$$B_\epsilon(T) = \begin{cases} g_2^{-1}\Big(g_1^{-1}\Big(\frac{b-\epsilon}{1-\epsilon}, 0\Big), \frac{b}{1-\epsilon}\Big), & \epsilon < \min(b, 1-b), \\ \infty, \quad \epsilon \geq \min(b, 1-b). \end{cases}$$

Lemma 4.3. *Consider T_b with jump function ρ_1.*

(i)

$$B_\epsilon(T_b) = G^{-1}(F_0^{-1})\Big(\frac{b}{1-\epsilon}\Big),$$

(ii)

$$\inf_{\epsilon<b<1-\epsilon} B_\epsilon(T_b) = \inf_{F_0^{-1}<t<\infty} G_t^{-1}\Big(2(1 - F_0(t)) + \frac{\epsilon}{1-\epsilon}\Big).$$

note. ここまでは、1989 の内容。

5. Fundamental Setting

1. $F_{H,\theta}(\nu)$ the distribution function of $|y - \theta^{\mathrm{T}} x|$

5.1. Notations.

$$T_\alpha(H) = \arg\min_{t \in \mathbb{R}^p} F_{H,t}^{-1}(\alpha)$$

5.2. lower bound.

The lower bound is given by

$$a(\epsilon) = \sup\{\|\theta\|; (1-\epsilon)F_{H_0,\theta}(\nu) + \epsilon \geq (1-\epsilon)F_{H_0,0}(\nu), \forall \nu \geq 0\}.$$

6. Main Results

Theorem 6.1. *Suppose that*

(i) $f_0(v) = F_0'(v)$ *is even and strictly unimodal.*
(ii) $P_{G_0}(\theta^{\mathrm{T}} x = 0) < 1, \forall \|\theta\| = 1.$
(iii) $G_0(x)$ *is elliptical.*

Then there exists α^* such that the least α^*-quantile estimate T_{α^*} is minimax bias in the class of residual admissible estimates.

note. 以下は、He and Simpson の結果

7. setting

7.1. Fisher consistent.

$\forall \theta \in \Theta$, $T(F_\theta) = \theta$.

7.2. locally linear.

Definition 7.1. Let $\{F_\theta; \theta \in \Theta\}$ be an L_1-differentiable family of distributions with finite Fisher information. A functional T is *locally linear* if it has the following properties.

(i) T has an influence function whose mean is 0 and finite variance.
(ii) T has the following expansion

$$T(F_{\theta+\delta}) - T(F_\theta) = \int \psi_\theta d(F_{\theta+\delta} - F_\theta) + o(\|\delta\|).$$

7.3. minmax.

Define

$$d_\nu(P, Q) = \sup|P(A) - Q(A)|$$

and

$$d_c(P, Q) = \inf\{\epsilon; Q = (1-\epsilon)P + \epsilon R \text{ for some distribution } R\}.$$

The variation gauge is

$$b_\nu(\epsilon, F_\theta) = \sup\{\rho(\theta, \eta); \eta \text{ such that } d_\nu(F_\theta, F_\eta) \leq \epsilon\}.$$

The contamination bias is

$$b_T(\epsilon, F_\theta) = \sup_{d_c(F_\theta, F) \le \epsilon} \rho(T(F), \theta).$$

Theorem 7.2.

$$\sup_{\eta; \rho(\theta,\eta) \le b_\nu(\epsilon/(1-\epsilon), F_\theta)} b_T(\epsilon; F_\eta) \ge \frac{1}{2} b_\nu\Big(\frac{\epsilon}{1-\epsilon}; F_\theta\Big)$$

note. つまり、Theorem 7.2 の右辺は、最大 bias の lower bound となる。これを達成する統計量は、min-max といえる。

8. Summary

8.1. **流れ-1.**

- Hosoya (1978) や Taniguchi (1981) は、真のスペクトルの ϵ contamination の場合について minimax を論じた。

- Martin, Yohai and Zamar (1989) は、母数の推定方式によって、minimax curve が異なることを明かした。

- Yohai and Zamar (1993) は、更に、quantile score による minimax curve は、admissible であることを明かした。

- Berrendero and Zamar (2001) は、説明変数が非対称分布、および定数項のある Regression の min-max について Bound を与えた。

- Berrendero, Mendes and Tyler (2007) は、Yohai and Zamar (1993) を否定した。quantile score は inadmissible となる理由は、回帰モデルでは、CM-estimate を構成することができる。ただし、CM-estimate は回帰モデルで構成できるが、ほかのモデルでは他の estimate を考える必要があり、論文では論ぜない。ただし、どんなモデルでも、quantile score は admissible とはならないだろうと予想。

8.2. **流れ-2.** こちらは、一般論。ϵ-contamination とは何かについて、分布間の距離で考える。

- Donoho and Liu (1988a) は、各距離による Robustness をまとめた。特に、L2 距離を Hermitian 距離と呼んで、良い性質を持つことを説いた。ここで、stability of variance 及び stability of quantity estimated について言及し、stability of variance は Huber の仕事で M-est を提案したのに対し、本論文は、後者について MD-est の重要性を説いた。

- Donoho and Liu (1988b) は、上記の良い性質を丁寧に説明。後に、depth や contour 等の概念へ発展。(ここから先は Follow できませんでした。)

- He and Simpson (1993) は、よりシャープな bound を与えた。更に、quantile score は、bound に到達することを示し、min-max 性を説いた。