UNIWERSYTET WARSZAWSKI WYDZIAŁ

MATEMATYKI, MECHANIKI I INFORMATYKI

KONRAD FURMAŃCZYK

ASYMPTOTYCZNE WŁASNOŚCI M-ESTYMATORÓW DLA DANYCH ZALEŻNYCH

AUTOREFERAT ROZPRAWY DOKTORSKIEJ

$\boxed{PLAN\ REFERATU}$

- Definicja M-estymatora
- Modele zależności
- Główne wyniki
- Metody dowodów
- Pomocnicze twierdzenia probabilistyczne

M-ESTYMATORY

Rozważmy ściśle stacjonarny ciąg zmiennych losowych $X_1,\ X_2, \ldots, X_n$ o wartościach w $\mathbb{R}^m$.

Realizacje tego ciągu będziemy nazywali obserwacjami. Rozkład brzegowy X_1, zależny od parametru θ, oznaczymy przez $\mathbb{P}_\theta$, gdzie $\theta \in \Theta \subset \mathbb{R}^k$. Zakładamy, że z rodziną rozkładów $\{\mathbb{P}_\theta\}$ możemy związać pewne mierzalne odwzorowanie $\psi : \Theta \times \mathbb{R}^m \to \mathbb{R}^k$ tak, że dla każdego $\theta \in \Theta$

$$\mathbb{E}_\theta \psi(\theta, X_1) = \int \psi(\theta, x) d\mathbb{P}_\theta(x) = 0\ .$$

DEFINICJA M-estymatorem parametru θ nazywamy każde mierzalne rozwiązanie θ_n równania

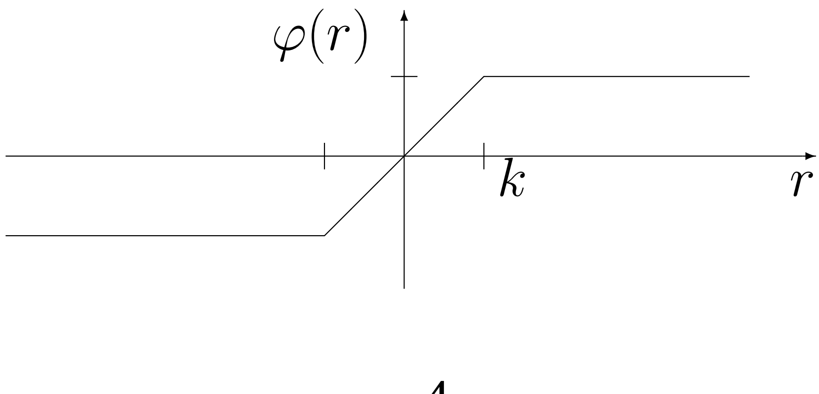
$$\Psi_n(\theta) := \frac{1}{n} \sum_{i=1}^{n} \psi(\theta, X_i) = 0\ .$$

PRZYKŁADY

- metoda najmniejszych kwadratów, wtedy $\psi(\theta, x) = x - \theta$, M-estymator $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$;

- metoda największej wiarogodności (dla niezależnych obserwacji), przy pewnych założeniach regularności $\psi(\theta, x) = \dfrac{\frac{\partial}{\partial\theta} f_\theta(x)}{f_\theta(x)}$, gdzie f_θ jest gęstością zmiennej losowej X_1;

- estymator Hubera (odporność), dla symetrycznych rozkładów prawdopodobieństwa $\psi(\theta, x) = \varphi(x - \theta)$,



MODELE ZALEŻNOŚCI

Głównym modelem zależności rozważanym w naszej pracy jest proces liniowy

(LP) $$X_j = \sum_{r=0}^{\infty} a_r Z_{j-r},$$

gdzie (Z_n) jest ciągiem zmiennych losowych i.i.d. oraz $\mathbf{E}(Z_1) = 0$, $\mathbf{E}(Z_1^2) = 1$, zaś (a_n) jest pewnym ciągiem liczbowym takim, że $\sum_{r=0}^{\infty} a_r^2 < \infty$. W pracy analizujemy tylko przypadek zależności bliskiego zasięgu (**ZBL**) procesu liniowego, tzn.

$$\sum_{j=0}^{\infty} |Cov(X_1, X_{1+j})| = \sum_{j=0}^{\infty} \left| \sum_{r=0}^{\infty} a_r a_{r+j} \right| < \infty \ .$$

W pracy rozpatrujemy warunek ($\mathbf{a}_0$) $\sum_{r=0}^{\infty} |a_r| < \infty$, który implikuje (**ZBL**). Rozpatrywane w pracy są również pokrewne modele zależności (**ZBL**): wielowymiarowy proces liniowy (**MLP**), obustronny proces liniowy (**LPO**), przekształcony proces gaussowski (**SGauss**).

GŁÓWNE WYNIKI

Wszystkie dalsze rozważania będą dotyczyły pewnego ustalonego (chociaż nieznanego) θ_0. Oznaczamy $\mathbb{P} = \mathbb{P}_{\theta_0}$ i $\mathbb{E} = \mathbb{E}_{\theta_0}$. Niech

$$\Psi(\theta) \;=\; \mathbb{E}\psi(\theta, X_1) \qquad \Psi_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \psi(\theta, X_i),$$
$$\Psi(\theta_0) \;=\; 0 \qquad\qquad \Psi_n(\theta_n) = 0.$$

ISTNIENIE M-ESTYMATORA

Rozważmy następujący warunek

(D) istnieje stała $c > 0$ oraz funkcja $M(x)$ takie że,

$$|D_j D_i \psi(\theta, x)| \leq M(x)$$

dla wszystkich $x \in \mathbb{R}$, wszystkich θ takich, że: $|\theta - \theta_0| < c$ oraz $\mathbb{E}\,|M(X_1)| < \infty$.

TWIERDZENIE (tw.3.11)

Założenia Ciąg obserwacji (X_n) jest ergodyczny, istnieje $(D\Psi(\theta_0))^{-1}$ oraz zachodzi warunek (**D**).

Teza Istnieje ciąg zmiennych losowych θ_n taki, że

$$\mathbb{P}(\Psi_n(\theta_n) = 0) \underset{n\to\infty}{\longrightarrow} 1$$

oraz

$$\mathbb{P}(\theta_n \to \theta_0) = 1.$$

REPREZENTACJE ASYMPTOTYCZNE

Podstawowym faktem ułatwiającym analizę M-estymatorów jest linearyzacja

$$\sqrt{n}(\theta_n - \theta_0) = -\mathbf{V}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi(\theta_0, X_i) + R_n ,$$

gdzie $\mathbf{V} = D\mathbb{E}\psi(\theta_0, X_1)$ lub$\mathbf{V} = \mathbb{E}D\psi(\theta_0, X_1)$.

Gdy reszta R_n jest postaci $R_n = o_{\mathbb{P}}(1)$, wtedy mówimy o reprezentacji Ghosha.

Reprezentacja Ghosha umożliwia udowodnienie asymptotycznej normalności **(AN)** M-estymatora, gdy
$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(\theta_0, X_i) \to^d \mathcal{N}_k(0, \mathbf{\Sigma})$ przy $n \to \infty$.

Wtedy oczywiście

$$\sqrt{n}(\theta_n - \theta_0) \to^d \mathcal{N}_k(0, \mathbf{V}^{-1}\mathbf{\Sigma}(\mathbf{V}^{-1})^T) \ .$$

Podamy teraz pewnien wariant jednego z głównych twierdzeń naszej pracy, tw. 3.48. Rozważmy następujące warunki:

(a$_0$**)** $\sum_{r=0}^{\infty} |a_r| < \infty$,

(b$_1$**)** Z_1 ma całkowalną funkcję charakterystyczną φ_Z, $\int |u| \, |\varphi_Z(u)| \, du < \infty$,

(b$_2(t)$**)** $\mathbb{E} \, |Z_1|^t < \infty$ dla pewnego $t \geq 2$.

(C) $\theta_n \to^{\mathbb{P}} \theta_0$,

(V) macierz $\mathbf{V} = D\,\mathbb{E}\psi(\theta_0, X_1)$ istnieje i jest nieosobliwa oraz $\mathbb{E}\,|\psi(\theta_0, X_1)|^2 < \infty$

TWIERDZENIE

Założenia: Θ jest zbiorem otwartym wymiaru k, zachodzą warunki: $(\mathbf{a}_0)$, $(\mathbf{b}_1)$, $(\mathbf{b}_2(k+1))$, **(C)**, **(V)**. Zachodzi jeden z warunków:

- $|\psi(\theta_1,\ x) - \psi(\theta_2,\ x)| \leq L(x)\,|\theta_1 - \theta_2|$ dla dowolnych $\theta_1, \theta_2 \in \Theta$ i dowolnego x istnieje funkcja $L \in \mathbb{L}^1(\mathbb{R})$ oraz $\|L(X_1)\|_{k+1} < \infty$,
- istnieje pewna stała C taka, że
$$\sup_{(\theta,x)\in\Theta\times\mathbb{R}} |D_\theta D_x \psi_i(\theta, x)| < C,$$

gdzie $\psi = (\psi_1, \psi_2, \ldots, \psi_k)$ oraz dodatkowo spełniony jest jeden z warunków:

- funkcja $\psi(\theta_0,\ \cdot)$ spełnia warunek Lipschitza,
- funkcja $\psi(\theta_0,\ \cdot) \in \mathbb{L}^1(\mathbb{R})$,

Teza Zachodzi reprezentacja Ghosha $(R_n = o_{\mathbb{P}}(1))$ i asymptotyczną normalność M-estymatorów

$$\sqrt{n}(\theta_n - \theta_0) \to^d \mathcal{N}_k \left(0, \mathbf{V}^{-1}\mathbf{\Sigma}(\mathbf{V}^{-1})^T\right),$$

gdzie elementy macierzy $\mathbf{\Sigma}$ są postaci

$$\begin{aligned}\sigma_{i,j} &= \mathbb{E}\left(\psi_i(\theta_0, X_1)\psi_j(\theta_0, X_1)\right) + \\ &+ 2\sum_{t=1}^{\infty} \mathbb{E}\left(\psi_i(\theta_0, X_1)\psi_j(\theta_0, X_{1+t})\right) .\end{aligned}$$

Przy pewnych założeniach na gładkość funkcji ψ otrzymujemy reprezentację Ghosha z resztą postaci $R_n = \mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}})$ (p. tw. 4.1). Spora część pracy poświęcona jest wyciąganiu wniosków z udowodnionej reprezentacji Ghosha w szczególnych przypadkach problemów M-estymacji takich jak estymacja parametru położenia (wn. 3.37) i estymacja współczynników regresji w modelach liniowych (tw. 3.56).

Wzmocnieniem reprezentacji Ghosha jest reprezentacja Bahadura, tzn. gdy reszta w linearyzacji estymatora jest postaci $R_n = \mathcal{O}_{p.n.}(\delta_n)$ dla pewnego ciągu $\delta_n \to 0$ przy $n \to \infty$.

Rozważmy następujący warunek:

($\mathbf{V}_1$) macierz $\mathbf{V} = \mathbb{E}D\psi(\theta_0, X_1)$ istnieje i jest nieosobliwa oraz $\mathbb{E}\left|\psi(\theta_0, X_1)\right|^2 < \infty$, TWIERDZENIE (wersja wn. 4.14)

Założenia Zachodzą warunki: **(a$_0$)**, $(\mathbf{b}_1)$, $(\mathbf{b}_2(Q))$ dla $Q > 2$, **(D)**, **(V$_1$)** oraz spełniony jest jeden z warunków:

- funkcje $\psi(\theta_0, \cdot)$ i $D\psi(\theta_0, \cdot)$ spełniają warunek Lipschitza,
- funkcje $\psi(\theta_0, \cdot)$, $D\psi(\theta_0, \cdot) \in \mathbb{L}^1(\mathbb{R})$.

Teza

$$\sqrt{n}(\theta_n - \theta_0) = -\mathbf{V}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi(\theta_0, X_i) + R_n,$$

gdzie $\mathbf{V} = \mathbb{E}D\psi(\theta_0, X_1)$ oraz dla dowolnego $\beta > 0$

$$R_n = \mathcal{O}_{p.n.}\left(\frac{(\log n)^{\frac{1}{Q}}(\log\log n)^{\frac{1+\beta}{Q}}}{\sqrt{n}}\right).$$

Ponadto otrzymujemy własność **(AN)**.

METODY DOWODÓW

Metoda Pollarda, odwołuje się do teorii procesów empirycznych. W powyższej metodzie oprócz kilku łatwych do sprawdzenia warunków występuje warunek asymptotycznej jednakowej ciągłości (**ASE**) procesu empirycznego

$$\upsilon_n\psi(\theta,\cdot)\ :=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\psi(\theta,X_i)-\mathbb{E}\psi(\theta,X_i)\}=\sqrt{n}(\Psi_n(\theta)-\Psi(\theta)).$$

(ASE) $\qquad \forall\,\varepsilon>0\,\forall\,\eta>0\,\exists\,\delta>0$

$$\limsup_{n\to\infty}\mathbb{P}\left(\sup_{|\theta-\theta_0|<\delta}|\upsilon_n\psi(\theta,\cdot)-\upsilon_n\psi(\theta_0,\cdot)|>\eta\right)<\varepsilon.$$

Warunek (**ASE**) otrzymamy z poniższego lematu, który jest wnioskiem twierdzenia Ledoux i Talagranda (91), wn. 11.8.

LEMAT (lemat 3.8)

Założenia Jeśli $dim(\Theta) = k$, proces empiryczny $\upsilon_n\psi(\theta, \cdot)$ posiada ciągłe trajektorie oraz dla pewnego $Q \geq 1$ spełnione są następujące warunki:

- $\|\upsilon_n\psi(\theta_1, \cdot) - \upsilon_n\psi(\theta_2, \cdot)\|_Q \leq C\,|\theta_1 - \theta_2|$ dla dowolnych $\theta_1, \theta_2 \in \Theta_\delta$ (z pewnego otoczenia θ_0), gdzie C jest pewną stałą;
- $\int_0^\delta \sqrt[Q]{N(\Theta_\delta, r)}dr < \infty$, gdzie $N(\Theta_\delta, r)$ oznacza minimalną liczbę kul o promieniach $\leq r$ pokrywających zbiór Θ_δ.

Teza

$$\mathbb{E} \sup_{|\theta-\theta_0|<\delta} |\upsilon_n\psi(\theta, \cdot) - \upsilon_n\psi(\theta_0, \cdot)| \leq$$

$$K \int_0^\delta \sqrt[Q]{N(\Theta_\delta, r)}dr \leq K' \int_0^\delta r^{-\frac{k}{Q}}dr.$$

Jak widzimy, żeby otrzymać warunek (**ASE**) wystarczy kontrolować przyrosty procesu empirycznego w normie $k+1-$szej (dla $Q=k+1$).

Metoda elementarna polega na rowinięciu funkcji ψ w szereg Taylora i prowadzi do silniejszej tezy, ale przy silniejszych założeniach na gładkość funkcji ψ. W ten sposób otrzymaliśmy reprezentację Bahadura (p. tw. 4.9) i reprezentację Ghosha z resztą postaci $\mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}})$ (p. tw. 4.1).

Pomocnicze twierdzenia probabilistyczne

Oszacowania momentów

TWIERDZENIE (tw. 6.2)

Założenia: $(\mathbf{a}_0)$, $(\mathbf{b}_1)$, $(\mathbf{b}_2(Q))$ dla $Q \geq 2$ oraz $\mathbb{E}\left|G(X_1)\right|^Q < \infty$. Jeśli spełniony jest jeden z warunków:

- funkcja G spełnia warunek Lipschitza,
- funkcja G jest całkowalna.

Teza

$$\mathbb{E}\left|\sum_{j=1}^{n}(G(X_j)-\mathbb{E}G(X_j))\right|^Q \leq C_* n^{\frac{Q}{2}},$$

gdzie $C_* = CC(G)$. W pierwszym przypadku

$$C(G)=(Lip(G))^Q, C=C_{Q,Z}\left(\sum_{r=0}^{\infty}|a_r|\right)^Q,$$

oraz w drugim przypadku

$$C(G)=\left(\|G(X_1)\|_Q+\left(\int|G(x)|\,dx\right)\right)^Q,$$

$C=C_{Q,Z}\left(\sum_{r=0}^{\infty}|a_r|\right)^Q$, gdzie stała $C_{Q,Z}$ zależy od rozkładu innowacji Z_1 i Q.

Pierwszy warunek występujący w Lemacie 3.8 otrzymujemy z Twierdzenia 6.2 podstawiając $G(x) = \psi(\theta_1, x) - \psi(\theta_2, x)$.

Centralne twierdzenie graniczne było nam potrzebne do wykazania asymptotycznej normalności M-estymatorów.

TWIERDZENIE (tw. 7.2)

Założenia: $(\mathbf{a}_0)$, $(\mathbf{b}_1)$, $(\mathbf{b}_2(2))$. Jeśli zachodzi jeden z warunków:

- funkcja G spełnia warunek Lipschitza,
- funkcja G jest całkowalna

oraz

$$\sigma^2 := \mathbb{E}\left(G(X_1)\right)^2 + 2\sum_{j=1}^{\infty} \mathbb{E}\left(G(X_1)G(X_{1+j})\right) > 0.$$

Teza

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(G(X_i) - \mathbb{E}G(X_i)\right) \to^d \mathcal{N}(0, \sigma^2)\,.$$

Wykazaliśmy również przy pewnych technicznych założeniach centralne twierdzenie graniczne dla ważonych sum $\sum_{i=1}^n b_i G(X_i)$ (p. tw. 7.10). Uzupełnieniem powyższych wyników było oszacowanie $p.n.$ z góry typu prawo iterowanego logarytmu (p. tw. 8.3) będące bezpośrednim wnioskiem z twierdzenia Moricza (76).

TWIERDZENIE

Założenia: $(\mathbf{a}_0)$, $(\mathbf{b}_1)$, $(\mathbf{b}_2(Q))$ dla $Q > 2$. Jeśli zachodzi jeden z warunków:

- funkcja G spełnia warunek Lipschitza,
- funkcja G jest całkowalna,

to dla dowolnego $\beta > 0$

$$\sum_{i=1}^{n} G(X_i) - \mathbb{E}G(X_i) = \mathcal{O}_{p.n.}(\sqrt{n}\delta_n),$$

gdzie $\delta_n = (\log \sqrt{n})^{\frac{1}{Q}} (\log \log \sqrt{n})^{\frac{1+\beta}{Q}}$.