目次

第 1 章

研究展望

ネタ帳

(1) 二重ロバスト推定と Double Machine Learning.

(2) 高次の影響関数への推定手法 (Robins が開拓) は因果推論ではとても有望視されている．Gateaux 微分ではだめでも，Malliavin 微分で理論を創れないか？これは robust 統計学の塗り替えに当たる試みになってしまう．

(3) Gaussian 近似．鈴木太慈先生と今泉先生の分野に近い．

1 で売れて，2 を温めるしかない．

Fisher 情報量は KL 情報量の二次形式．基本的に解析学と同じ発展経路をたどるんだろうな．

頑健性が反脆弱性になったら機械学習なんじゃないのか？ミスから学ぶ．

影響関数が存在するようなパラメータ付けを正則というらしい．となると統計モデルの解析学はここからじゃないと始まらないじゃないか！

1.1　頑健性についてのノート

用語法 1.1.1. 古典的な統計手法は，仮定された正規分布よりも尾が重い標本分布に対してとても感度が高く，すなわち外れ値に対して脆弱である．これを distributional distortion に対する distributional robustness という．一部では resistant 統計量ともいう．そして，robust の語を，モデル設定や推定量についての仮定への違反に対して使うために控える．だが，通常 robustness といったとき，distributional robustness を指す．外れ値を，経験分布とデルタ分布との凸結合によって表して，これの Gateaux 片側微分係数によって定まる影響関数で，頑健性の定義とするのがひとつの流儀である．これを B-ロバスト統計量とよんでいる．

(1) B-ロバスト統計量は，バイアスについて頑健である．

例 1.1.2. 確率分布の中心的傾向 (central tendency) の指標として，算術平均は脆弱であるが，中央値は頑健である．M-推定量は普通頑健な方である．

1.2　Locally Robust Semiparametric Estimation

1st step に対して頑健な GMM 推定量を定める局所頑健＝直交モーメント関数の標準的構成法．1st step というのは，モデル選択や正則化が産むバイアスを含むので，特に機械学習に応用先がある．

1.2.1　疑問点

(1) Many economic and causal parameters depend on nonparametric or high dimensional first steps. 全般的に"first step"が何を指しているかわからない．

1.2.2 Introduction

この論文では，一般化モーメント法の推定関数＝モーメント関数を，局所頑健＝直交に構成する方法を示す．これが定める一般化モーメント推定量を，ここでは脱偏済み推定量 (debiased estimator) という．

> We show that such moment functions can be constructed by adding to identifying moment functions the non-parametric influence function from the effect of the first step on identifying moments.

1.2.3 Debiased GMM estimator

記法 1.2.1. $\theta \in \Theta \subset \mathbb{R}^p$ を母数，$\gamma : \Xi \to \mathbb{R}$ を未知関数（局外母数），$W : \Omega \to \mathcal{X}$ を観察されたデータ，$w \in \mathcal{X}$ をその実現値とする．真値 $\theta_0 \in \Theta, \gamma_0 \in \Xi \times \mathbb{R}$ に関して $E[g(W, \gamma_0, \theta_0)] = 0$ を満たす $g :\to \mathbb{R}^q$ を所与の関数とする．$\theta_0 \in \Theta$ は $E[g(W, \gamma_0, \theta)] = 0$ を満たす唯一の値であるという識別性条件（すなわち，θ_0 を，一般化モーメント $E[g(W, \gamma_0, \theta)]$ の値で識別できる．このことを"**identifying moment function** g"と呼んでいる）を仮定する．

議論 1.2.2. まず，$\widehat{\gamma}$ を γ_0 の first-step 推定量としよう．すると，plug-in method として，観測値 w_i と推定値 $\widehat{\gamma}$ を代入して $g(w_i, \gamma, \theta)$ を推定関数とし，経験測度について平均を取ると，標本モーメントの推定量 $\frac{\sum_{i=1}^n g(w_i, \widehat{\gamma}, \theta)}{n}$ を得る．これの特定のノルム (論文では quadratic form と表現されてる) を最小にする θ を，"plug-in" GMM estimator とすれば良い．しかし，この推定量は，最初の $\widehat{\gamma}$ を得たときのモデル選択，ひいては正則化法に多大な影響を受ける．[†1]

1.2.3.1 直交モーメント関数への準備

F をモデルの CDF とし，F_0 を真の累積分布関数とする．推定量 $\widehat{\gamma}$ に対して，F の関数 $\gamma(F)$ が存在して，$F = F_0$ のときこれは推定量 $\widehat{\gamma}$ の確率収束極限であるとする．

H を別の分布関数とし，凸結合 $F_\tau := (1 - \tau)F_0 + \tau H$ を考える．絶対連続性など理想的な状況が揃い，$E[g(W, \gamma(F_\tau), \theta)]$ の $\tau = 0$ における右 Gateaux 微分係数の微分 ϕ を影響関数という：

$$\frac{d}{d\tau} E[g(W, \gamma(F_\tau), \theta)] = \int \phi(w, \gamma_0, \alpha_0, \theta) H(dw)$$

どういう統計量 T についての影響関数かというと，一般化モーメント $\mu(F) := E[g(W, \gamma(F), \theta)]$ に関する影響関数である．

この影響関数 $\phi(w, \gamma, \alpha, \theta)$ を**ノンパラメトリック影響関数**といい，最初の推定量 γ が，標本モーメントの推定値 $\mu(F)$ に与える局所的な影響を記述しているとみなせる．

ただし，$E[\phi(W, \gamma_0, \alpha_0, \theta)] = 0$ を満たすとする．

定義 1.2.3 (orthogonal moment function)**.** 識別可能性を持つ一般化モーメント関数 (identifying moment function)g と，ノンパラメトリック影響関数（の微分）ϕ の和

$$\psi(W, \gamma, \alpha, \theta) := g(W, \gamma, \theta) + \phi(W, \gamma, \alpha, \theta)$$

を**直交モーメント関数**という．

要諦 1.2.4. ϕ が，γ の影響を打ち消すから，ψ は頑健になる，ということか！条件 $E[\phi(W, \gamma_0, \alpha_0, \theta)] = 0$ により，任意の経験測度について，$\widehat{\phi}_n \xrightarrow{p} 0$ が成り立つ．よって，ϕ の影響はこの意味ではなく，ただ $\widehat{\gamma}_l$ の first-order effect を打ち消すだけの役割を果たす．また $\widehat{\phi}$ の構成も，モデルの仮定に依らず，関数解析の言葉でノンパラメトリックに指定している．

定理 1.2.5.

(1) γ, α の推定量による ψ の期待値への影響の 1 次の項は消える．

(2) θ の推定量による ϕ の期待値への影響の 1 次の項は消える．

[†1] 星野先生の本では，first-step は無限次元ではなく，傾向スコアに関するパラメトリックモデルだとして並行な議論をしている．時代は進んだ．

1.2.3.2　cross-fitting

"plug-in" method と違って，θ も第一段階で推定してしまうのか．
そして，このようなことをする理由は，平均を取るところの標本の選び方からバイアスを抜くため．ある種のリサンプリング法？

この ψ を用いて，debiased GMM estimator を構成するのだが，ここで **cross-fitting** を用いる．これは sample splitting の一種である．標本の添字集合 $[n]$ を L 個に分割し，$\widehat{\gamma}_l, \widehat{\alpha}, \widehat{\theta}_l$ $(l \in [L])$ を，成分 $I_l \subset [n]$ に属する標本を使わずに算出した推定量とする．そして，脱偏済み標本モーメント関数 $\widehat{\psi}$ を

$$\widehat{g}(\theta) = \frac{1}{n}\sum_{l=1}^{L}\sum_{i\in I_l} g(W_i, \widehat{\gamma}_l, \theta), \quad \widehat{\phi} = \frac{1}{n}\sum_{l=1}^{L}\sum_{i\in I_l} \phi(W_i, \widehat{\gamma}_l, \widehat{\alpha}_l, \widehat{\theta}_l)$$

について

$$\widehat{\psi}(\theta) = \widehat{g}(\theta) + \widehat{\phi}$$

と定め，特定の正な半正定値行列 $\widehat{W}$ について，$\widehat{\theta} := \operatorname*{arg\,min}_{\theta\in\Theta} \widehat{\psi}(\theta)^\top \widehat{W}\widehat{\psi}(\theta)$ として debiased GMM estimator $\widehat{\theta}$ を定義すれば良い．

注 1.2.6 (θ の第一段階推定量について)**.** $\widehat{\theta}_l$ についてだが，通常の GMM として

$$\widehat{\theta}_l := \operatorname*{arg\,min}_{\theta\in\Theta} \widehat{g}_l(\theta)^\top \widehat{\Upsilon}_l \widehat{g}_l(\theta), \quad \widehat{g}_l(\theta) = \frac{1}{n-n_l}\sum_{l'\neq l}\sum_{i\in I_{l'}} g(W_i, \widehat{\gamma}_{l'}, \theta)$$

] と定めても良い．ただし，$\widehat{\Upsilon}_l$ の構成には I_l に入っていない標本のみを使い，cross-fitting を取り入れる．このあとに debiased GMM を計算するセットを，何度か繰り返して最終的な debiased GMM としても良い（本当か？）

1.2.3.3　**効率**

こうして得た debiased GMM の効率は，以下の 3 つの要因のみのよる．

(1) モーメント関数 g の選び方
(2) 第一段階推定量 $\widehat{\gamma}, \widehat{\alpha}, \widehat{\theta}$
(3) 重み付け行列 $\widehat{W}$

(3) については別の議論で済んでいて，標準的な取り方 Ψ がある．

1.2.3.4　**例 1：条件付き共分散**

例 1.2.7. $W = (X, Y, Z)$ を観測データ，$\alpha_0(X) := E[Z|X]$ を共変量と割当の関係，$\gamma_0(X) := E[Y|X]$ を共変量と平均処置効果の関係とする．モーメント関数を $g := \alpha_0 \otimes \gamma_0$ と定めて，$\theta_0 = E[Z\gamma_0(X)] = E[\alpha_0(X)\gamma_0(X)] = E[E[Z|X]\cdot E[Y|X]]$ と仮定する．このような設定は，条件付き共分散の平均 $E[\mathrm{Cov}(Z, Y|X)] = E[ZY] - \theta_0$ などに登場する．

このとき，

$$g(w, \gamma, \theta) := z\gamma(x) - \theta, \quad \phi(w, \gamma, \alpha) = \alpha(x)[y - \gamma(x)]$$

と定めると，

1.2.3.5　例 2：条件付き分位点の関数

1.2.4　例 3：

1.2.5　Neyman 直交性

Gateaux 微分における展開で一次項が消えていることをいう．どういう内積を仮定しているんだ？

定義 1.2.8. 未知関数 γ,α が，モーメント $\overline{\psi}(\gamma,\alpha,\theta):=E[\psi(W,\gamma,\alpha,\theta)]$ に 1 次の影響を及ぼさないことを，**Neyman 直交**するという．すなわち，F を変数 W の累積分布関数とするとき，$\widehat{\alpha}\xrightarrow{p}\alpha(F)$ と表すと，影響関数 $\phi(w,\gamma,\alpha,\theta)$ の識別可能性条件より，$0=E_F[\phi(W,\gamma(F),\alpha(F),\theta)]$ が成り立つ．$F_\tau=(1-\tau)F_0+\tau H$ を代入して $\tau=0$ について微分することで，

$$\begin{aligned}0&=\frac{\partial}{\partial\tau}\int\phi(w,\gamma(F_\tau),\alpha(F_\tau),\theta)(-F_0+H)(dw)\\&=\int\phi(w,\gamma_0,\alpha_0,\theta)H(dw)+\frac{\partial}{\partial\tau}E[\phi(W,\gamma(F_\tau),\alpha(F_\tau),\theta)]\\&=\frac{\partial}{\partial\tau}E[g(W,\gamma(F_\tau),\theta)]+\frac{\partial}{\partial\tau}E[\phi(W,\gamma(F_\tau),\alpha(F_\tau),\theta)]=\frac{\partial}{\partial\tau}\overline{\psi}(\gamma(F_\tau),\alpha(F_\tau),\theta)\end{aligned}$$

を得て，たしかに 1 次の微分係数がきえている．これは $(\gamma(F_\tau),\alpha(F_\tau))$ に関する方向微分だと思えるはず．

定理 1.2.9. 次の 3 条件が成り立つとき，Neyman 直交性を満たす．

(1) ノンパラメトリック影響関数の識別可能性（とは少しずれるらしい）：$E[\phi(W,\gamma_0,\alpha_0,\theta)]=0$.
(2) 影響関数の微分可能性：ある $\overline{\tau}>0$ が存在して，任意の $\tau\in[0,\overline{\tau})$ について $\displaystyle\int\phi(w,\gamma(F_\tau),\alpha(F_\tau),\theta)F_\tau(dw)=0$.
(3) 連続性：$\displaystyle\int\phi(w,\gamma(F_\tau),\alpha(F_\tau),\theta)F_0(dw)$, $\displaystyle\int\phi(W,\gamma(F_\tau),\alpha(F_\tau),\theta)H(dw)$ はいずれも $\tau=0$ で連続.

1.2.5.1　二重ロバスト性へむけて

定理 1.2.10 (α の十分条件). α が次の 2 条件を満たすならば，$E[\phi(W,\gamma_0,\alpha,\theta)]=0$. よって，$\overline{\psi}=E[\psi(W,\gamma_0,\alpha,\theta)]=0$.

(1) F_α が存在して，$\alpha(F_\alpha)=\alpha$ かつ $\exists_{\overline{\tau}>0}\ \forall_{t\in\overline{\tau}}\ \gamma(F_\tau^\alpha)=\gamma_0$.
(2) $\displaystyle\frac{d}{d\tau}\int g(w,\gamma(F_\tau^\alpha),\theta)F_\alpha(dw)=\int\phi(w,\gamma_0,\alpha,\theta)F_0(dw)$.

定理 1.2.11 (γ の十分条件). γ のノルム $\|-\|$ と，γ_0 を含む集合 Γ と，累積分布関数の集合 $\mathcal{H}$ は次の 4 条件を満たすならば，$\forall_{\delta\in\Gamma}\ \overline{\psi}_\gamma(\delta,\alpha_0,\theta_0)=0$

(1) $\alpha(F_\tau)=\alpha_0$ で，Neyman 直交性を満たす.
(2) $\overline{\psi}(\gamma,\alpha_0,\theta_0)$ は γ_0 にて，Γ に近接して Hadamard 微分可能で，Γ 上で定義された偏導関数 $\overline{\psi_\gamma}(\delta,\alpha_0,\theta_0)$ を持つ.
(3) $\gamma(F_\tau)$ は $\tau=0$ にて Hadamard 微分可能.
(4) $\displaystyle\left\{\frac{\partial\gamma(F_\tau)}{\partial\tau}\right\}_{H\in\mathcal{H}}$ の閉包は Γ に等しい.

また，$\overline{\psi}(\gamma,\alpha_0,\theta_0)$ は γ_0 の近傍で二階連続微分可能で，近傍にて $\exists_{C\in\mathbb{R}}\ \|\overline{\psi}(\gamma,\alpha_0,\theta_0)\|\le C\|\gamma-\gamma_0\|^2$.

1.2.6　Plug-in GMM との比較

1.2.7　α_0 の自動推定

1.2.8　二重ロバスト性

ノンパラメトリック影響関数を，identifying モーメント関数 g に加える，という手法で，二重ロバスト性を満たす ψ を構成できる条件を与える．

議論 1.2.12. 一般に，局外母数 γ を正しく推定することは困難なので，二重ロバスト性はもはや必須となりつつある．これは，モーメント関数 g に関するクラスで，ある二重ロバスト条件を満たすことをいう．いままでノンパラメトリック影響関数の mean zero condition は $E[\phi(W,\gamma_0,\alpha_0,\theta)]=0$ としてきたが，$E[\phi(W,\gamma,\alpha_0,\theta)]=E[\phi(W,\gamma_0,\alpha,\theta)]=0$ と強める．

1.2.8.1　二重ロバスト性の特徴付け

定義 1.2.13. Γ を γ の第一段階推定量のありえる値域となる位相線型空間とする．

$$\forall_{\gamma\in\Gamma}\ \forall_{\alpha,\theta}\quad 0=\overline{\psi}(\gamma,\alpha_0,\theta_0)=\overline{\psi}(\gamma_0,\alpha,\theta)$$

を満たすとき，モーメント $\overline{\psi}$ は二重に頑健であるという．

系 1.2.14. Γ を線型空間とする．$\overline{\psi}(\gamma,\alpha_0,\theta_0)$ が γ について線型で，定理 31.2.11 の条件が成り立つとき，二重に頑健である．

［証明］. 定理 31.2.11 の意味での Γ は γ の近傍であり，局所的な集合であったが，これが生成する線型空間の全体にまで，性質 $\overline{\psi}(\gamma,\alpha_0,\theta_0)=0$ が延長できるため．■

定理 1.2.15 (Gateaux 微分のことばで述べ直す)**.** Γ を線型空間とする．次の 2 条件は同値．

(1) $\psi(w,\gamma,\alpha,\theta)$ は二重に頑健である．
(2) $\overline{\psi}(\gamma,\alpha_0,\theta_0)$ は γ について affine で，$\forall_{\gamma\in\Gamma}\ \dfrac{\partial\overline{\psi}((1-\tau)\gamma_0+\tau\gamma,\alpha_0,\theta_0)}{\partial\tau}|_{\tau=0}=0.$

系 1.2.16. $g(W,\gamma,\theta_0)$ と $\phi(W,\gamma,\alpha_0,\theta_0)$ が γ について線型であるとき，これが定める $\overline{\psi}(\gamma,\alpha,\theta)$ は二重に頑健である．

1.2.8.2　第一推定量の条件付きモーメント条件下での二重頑健推定量の構成法

議論 1.2.17. 第一段階推定量 γ_0 は，ある線型汎関数 $\lambda(W,-):\Gamma\to\mathbb{R}$ について，$E[\lambda(W,\gamma_0)|X]=0$ を満たすとする．X は共変量や操作変数を想定すれば良い．

この下で 2SLS(two stage least squares) 推定量 $\widehat{\gamma}$ を考えると，$\gamma(F):=\operatorname*{arg\,min}_{\gamma\in\Gamma}E_F[(E_F[\lambda(W,\gamma)|X])^2]$ として，$\phi(w,\gamma,\alpha,\theta)$ が存在するとき，ある $\alpha(x,\theta)$ について，

$$\phi(w,\gamma,\alpha,\theta)=\alpha(x,\theta)\lambda(W,\gamma)$$

が成り立つ．特に，ϕ は γ について線型である．よって，ψ の線形性は g の線形性による．

定理 1.2.18. 次の 2 条件は同値．

(1) $\psi(W,\gamma,\alpha,\theta)=g(W,\gamma,\theta)+\alpha(X)\lambda(W,\gamma)$ は二重に頑健である．
(2) $\forall_{\gamma\in\Gamma}\ E[g(W,\gamma,\theta_0)]=-E[\alpha_0(X)\lambda(W,\gamma)]$.

要諦 1.2.19. 有効スコアの場合に似ている関数形をもっているらしい．

例 1.2.20. γ を回帰関数，$\lambda(W,\gamma)=Y-\gamma(X)$ を誤差を表す線型汎関数，$m(w,\gamma)$ を線型汎関数とし，$\theta_0=E[m(W,\gamma_0)]$ の推定を目指す．モーメント関数を $g(w,\gamma,\theta):=m(w,\gamma)-\theta$ とおくと，これは識別可能な上に γ について線型で，λ も affine だから，定理 8 の前提条件を満たす．よって，

系 1.2.21. $\alpha_0(x) \in \mathcal{L}^2(\mathcal{X})$ が存在し，$\forall_{\gamma(X)\in\mathcal{L}^2(\mathcal{X})}\ E[m(W,\gamma)] = E[\alpha_0(X)\gamma(X)]$ ならば，$\psi(w,\gamma,\alpha,\theta) = m(w,\gamma)-\theta+\alpha(x)[y-\gamma(x)]$ は二重に頑健である．

1.2.8.3　局外母数が共変量の確率密度関数であるときの二重頑健推定量の構成

1.2.8.4　識別可能性

定理 1.2.22.

1.2.8.5　plug-in GMM 推定量の部分的頑健性

定義 1.2.23. モーメント関数 g が**部分的に頑健**であるとは，$\exists_{\overline{\gamma}\neq\gamma_0}\ E[g(W,\theta_0,\overline{\gamma})] = 0$ がなりたつことをいう．

1.2.9　漸近理論

1.3　Characterization of parameters with a mixed bias property

1.3.1　Introduction

記法 1.3.1. P に独立に従う確率ベクトル $O:\Omega\to\mathcal{D}$ の列 $O^n:\Omega\to\mathcal{D}^n$ を考える．モデルを $\mathcal{M} := (P_\eta)_{\eta\in\eta}$ とし，汎関数 $\chi(\eta)\in\mathbb{R}$ の推定を試みる．

議論 1.3.2 (debiased-GMM の動機：ワンステップ推定量)**.** 前の論文のように，plug-in 推定量 $\chi(\widehat{\eta})$ は $\sqrt{n}$-一致性を持たない：$\chi(\widehat{\eta})-\chi(\eta)\neq O_p(1/\sqrt{n})$．よって，漸近有効性が足りないから，バイアスを減らす余地がある．一つの解決法は，ワンステップ推定量（のセミパラメトリック化）である．χ_η^1 をうまく選ぶことで，$\widehat{\chi} := \chi(\widetilde{\eta}) + \mathbb{P}_n\chi_{\widehat{\eta}}^1$ をワンステップ推定量とする．χ_η^1 の上手な選び方は，前の論文のように，影響関数とすることである．

議論 1.3.3 (影響関数を採用する動機)**.** 正規分布に収束させたい項 $\sqrt{n}(\widehat{\chi}-\chi(\eta))$ は

$$\sqrt{n}(\widehat{\chi}-\chi(\eta)) = \sqrt{n}(\chi(\widehat{\eta})-\chi(\eta)+E_\eta(\chi_{\widehat{\eta}}^1)) + \mathbb{G}_n(\chi_{\widehat{\eta}}^1-\chi_\eta^1) + \mathbb{G}(\chi_\eta^1)$$

と展開できる．

(a) 第 3 項 $\mathbb{G}_n(\chi_{\widehat{\eta}}^1) = \sqrt{n}\mathbb{P}_n(\chi_\eta^1 - E_\eta(\chi_{\widehat{\eta}}^1))$ は正規分布に収束する．
(b) 第 2 項 $\mathbb{G}_n(\chi_{\widehat{\eta}}^1-\chi_\eta^1)$ は $o_p(1)$ である．一般にモデル $\mathcal{M}$ が大きくないとき (Donsker であるとき) はこれが示せ，一般の場合も cross-fitting によりこれが可能になる．$\mathcal{D}_n$ を複数に分割し，一部で推定量を構成し，残った部分でワンステップ推定量を構成する．これを繰り返し，全部の平均を最終的な推定量とする．
(c) 第 1 項が $o_p(1)$ ならば，$\sqrt{n}(\widehat{\chi}-\chi(\eta))$ は正規分布に確率収束する．

すなわち，$\chi(\widehat{\eta})-\chi(\eta)-E_\eta(\chi_{\widehat{\eta}}^1) = o_p(n^{-1/2})$ が満たされれば良い．$E_\eta(\chi_{\widehat{\eta}}^1)$ が，$\chi(\eta)$ の方向 $\widehat{\eta}-\eta$ への微分係数となっていれば良いわけである．

影響関数が存在するようなパラメータ付けを正則という．

1.3.1.1　枠組み

> 因果推論における指数型分布族？
> この mixed bias property ということばで，いままでの 2 つのクラス（Robins et al 2008 と Chernozhukov et al 2018）をまとめ上げた．モーメント関数の識別可能性条件などは，自動的に満たされる．

記法 1.3.4. $\mathcal{M}$ はノンパラメトリックとする：任意の $P\in\mathcal{M}$ におけるパラメトリックなサブモデルのスコアが生成する閉線形空間は $L_2(P)$ に等しい．$\chi(\eta)$ を正則なパラメータ付けとし，χ_η^1 をその一意な影響関数とする．次の mixed bias property を満たすとする．

定義 1.3.5 (mixed bias property)**.** 任意のパラメータ $\eta\in\eta$ について，関数 $a(Z) := a(Z;\eta)$ と $b(Z) := b(Z;\eta)$ が存在し，また η

に依らない関数 $S_{ab} := s_{ab}(O) : \mathscr{D} \to \mathbb{R}; o \mapsto s_{ab}(o)$ が存在して，

$$\forall_{\eta' \in \eta} \quad \chi(\eta') - \chi(\eta) + E_\eta(\chi^1_{\eta'}) = E_\eta[S_{ab}(a'(Z) - a(Z))(b'(Z) - b(Z))]$$

が成り立つ．ただし，$a'(Z) = a(Z; \eta'), b'(Z) = b(Z; \eta')$ と表した．

要諦 1.3.6. これは，$\chi(\eta) + \chi^1_\eta$ は，a, b という 2 つの関数を通じてのみ η に依存しており，したがってワンステップ推定量 $\widehat{\chi(\eta)}$ は推定量 $\widehat{a}, \widehat{b}$ を通じてのみ $\widehat{\eta}$ に依存する．

このことにより，$\int (\widehat{a}(z) - a(z))^2 dP_\eta(z) = O_p(\gamma_{a,n})$ かつ $\int (\widehat{b}(z) - b(z))^2 dP_\eta(z) = O_p(\gamma_{b,n})$ を満たすとき，$\chi(\widehat{\eta}) - \chi(\eta) - E_\eta(\chi^1_{\widehat{\eta}}) = O_p(\gamma_{a,n}\gamma_{b,n})$ が成り立つ．だから，cross-fitting が採用されたならば，$\widetilde{\chi}$ は **rate double robustness property** を持つ．すなわち，$\gamma_{a,n} = o(1), \gamma_{b,n} = o(1), \gamma_{a,n}\gamma_{b,n} = o(n^{-1/2})$ が成り立つならば，$\sqrt{n}(\widetilde{\chi} - \chi(\eta))$ は平均 0 の正規分布に収束する．

1.3.2　mixed bias property を持った影響関数の特徴付け

1.3.3　局外関数の特徴付け

1.3.4　例

1.3.5　Final remarks

1.4　De-Biased Machine Learning of Global and Local Parameters Using Regularized Riesz Representers

1.5　Double Debiased Machine Learning for Treatment and Structural Parameters

1.6　ON GAUSSIAN APPROXIMATION FOR M-ESTIMATOR

1.7　GAUSSIAN APPROXIMATION OF SUPREMA OF EMPIRICAL PROCESSES

参考文献

[1] Victor Chernozhukov, Juan Carlos Escanciano, Hidehiko Ichimura, Whitney K. Newey, James M. Robins - Locally Robust Semiparametric Estimation

[2] Andrea Rotnitzky, Ezequiel Smucler, James M. Robins - Characterization of parameters with a mixed bias property
生物統計分野の研究者が，二重ロバスト推定が有効ないままでにない関数クラスを発見した．これの必要十分条件を特定できないだろうか？

[3] De-Biased Machine Learning of Global and Local Parameters Using Regularized Riesz Representers

[4] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, James Robins - Double Debiased Machine Learning for Treatment and Structural Parameters

[5] Masaaki Imaizumi, Taisuke Otsu - ON GAUSSIAN APPROXIMATION FOR M-ESTIMATOR

[6] Victor Chernozhukov, Denis Chetverikov, Kengo Kato - GAUSSIAN APPROXIMATION OF SUPREMA OF EMPIRICAL PROCESSES