Semiparametric Models

Mona ちゃん
ケロロ大学タママ学部ペコポン研究所

December 17, 2024

1. 部分線形モデル
2. MINPIN 推定量
3. セミパラメトリック効率性限界
4. Appendix: 関数解析の基礎

1. 部分線形モデル
2. MINPIN 推定量
3. セミパラメトリック効率性限界
4. Appendix: 関数解析の基礎

ノンパラモデルにおける「次元の呪い」

- ノンパラメトリックモデルは、関数形の特定化に関する仮定が一番少ないため、特定化に関する誤りを犯す可能性は低い。
- しかし、その代償として、説明変数の数が増えるにつれて推定の精度が急速に下がっていく（**「次元の呪い」**）という問題点がある。
- ノンパラモデルはパラメトリックモデルと比べて結果の解釈が難しい。

パラメトリック構造の導入による次元の縮小

- パラメトリックな構造を導入することで、ノンパラパートにおける説明変数の次元を縮約し、ノンパラモデルの欠点を解消しようとしたものが**セミパラメトリックモデル**。
- セミパラモデルは、パラメトリックモデルよりも柔軟性があり、かつノンパラパートにおける説明変数の数を抑えることで「次元の呪い」を回避している。

部分線形モデル

- **部分線形モデル**は以下のように与えられる。

$$Y_i = X_i'\beta + g(Z_i) + u_i$$

$$E[u_i|X_i, Z_i] = 0$$

$$E[u_i|X_i, Z_i] = \sigma^2(X_i, Z_i)$$

  X_i は p 次元、Z_i は q 次元とする。
- $g()$ の関数形は特定化しない。
- パラメトリックパートの β とノンパラパートの $g()$ の推定が目的。

識別のための条件

- パラメータ β が切片を含む場合 (=X_i が定数項を含む場合)、β は識別不能となる。
- 理由:
  定数項 α_1、関数 g_1、0 でない定数 c について

$$\alpha_2 = \alpha_1 + c$$

$$g_2 = g_1 - c$$

  と定義すると、$\alpha_1 + g_1 = \alpha_2 + g_2$ となるため、α_1 と α_2、g_1 と g_2 が識別できなくなる。

Robinsonの推定方法1

- モデルの式

$$Y_i = X_i'\beta + g(Z_i) + u_i$$

の両辺の Z_i で条件付けた条件付き期待値を取ると

$$E[Y_i|Z_i] = E[X_i|Z_i]'\beta + g(Z_i)$$

となる。上の２式の差を取ると、

$$Y_i - E[Y_i|Z_i] = (X_i - E[X_i|Z_i])'\beta + u_i$$

ノンパラパートが消えた!

Robinsonの推定方法2

- $g()$ の消滅と同時に未知なるもの $E[Y_i|Z_i], E[X_i|Z_i]$ が登場したが、ひとまず無視して

$$\tilde{Y}_i = Y_i - E[Y_i|Z_i]$$

$$\tilde{X}_i = X_i - E[X_i|Z_i]$$

と新しい変数 $\tilde{Y}_i, \tilde{X}_i$ を定義する。すると...

$$\tilde{Y}_i = \tilde{X}_i'\beta + u_i$$

これは線形回帰モデル!

Robinsonの推定方法3

- $\tilde{Y}_i, \tilde{X}_i$ に隠した $E[Y_i|Z_i], E[X_i|Z_i]$ に目をつむれば、OLS により「実行不可能 (infeasible) な」推定量 $\hat{\beta}_{\text{inf}}$

$$\hat{\beta}_{\text{inf}} = \left(\sum_{i=1}^{n} \tilde{X}_i \tilde{X}_i{}' \right)^{-1} \sum_{i=1}^{n} \tilde{X}_i \tilde{Y}_i$$

が得られる。

$\hat{\beta}_{\text{inf}}$ の漸近正規性 1

- $\hat{\beta}_{\text{inf}}$ に $\tilde{Y}_i = \tilde{X}_i'\beta + u_i$ を代入し、整理すると、

$$\sqrt{n}(\hat{\beta}_{\text{inf}} - \beta) = \left(\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_i\tilde{X}_i{}'\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{X}_i\tilde{u}_i$$

- 大数の法則より、

$$\frac{1}{n}\sum_{i=1}^{n}\tilde{X}_i\tilde{X}_i{}' \xrightarrow{p} E[\tilde{X}_i\tilde{X}_i{}']$$

$\Phi = E[\tilde{X}_i\tilde{X}_i{}']$ と置く。

$\hat{\beta}_{\text{inf}}$ の漸近正規性 2

- 中心極限定理より、

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{X}_i\tilde{u}_i \xrightarrow{d} N(0, V[\tilde{X}_i\tilde{u}_i])$$

- $\Psi = V[\tilde{X}_i\tilde{u}_i] = E[\sigma^2(X_i, Z_i)\tilde{X}_i\tilde{X}_i{}']$ と置くと、

$$\sqrt{n}(\hat{\beta}_{\text{inf}} - \beta) \xrightarrow{d} \Phi^{-1}N(0, \Psi) = N(0, \Phi^{-1}\Psi\Phi^{-1})$$

$E[Y_i|Z_i], E[X_i|Z_i]$ と向き合う

- これまでわかっているものとしてきた $E[Y_i|Z_i], E[X_i|Z_i]$ をノンパラメトリックに推定しよう

$$\hat{Y}_i = \hat{E}[Y_i|Z_i] = \frac{\sum_{i=1}^{n} Y_j K(\frac{Z_i - Z_j}{h})}{\sum_{i=1}^{n} K(\frac{Z_i - Z_j}{h})}$$

$$\hat{X}_i = \hat{E}[X_i|Z_i] = \frac{\sum_{i=1}^{n} X_j K(\frac{Z_i - Z_j}{h})}{\sum_{i=1}^{n} K(\frac{Z_i - Z_j}{h})}$$

- これらをプラグインすれば「実行可能 (feasible) な」β の推定量が得られる。

ランダムな分母問題

- $E[Y_i|Z_i], E[X_i|Z_i]$ のノンパラ推定量 $\hat{Y}_i, \hat{X}_i$ の分母は $f(Z_i)$ のノンパラ推定量 $\hat{f}(Z_i)$ である。
- ある点 z で $f(z)$ が 0 に近い値を取る場合、$\hat{f}(z)$ も 0 に近い値を取るため $\hat{Y}_i, \hat{X}_i$ の推定精度が悪くなってしまう。

トリミング

- そこで指示関数 $\mathbf{1}_i = \mathbf{1}(\hat{f}(Z_i) \geq b)$ を用いて以下のような推定量を考える。

$$\hat{\beta} = \left\{ \sum_{i=1}^{n} (X_i - \tilde{X}_i)(X_i - \tilde{X}_i)' \mathbf{1}_i \right\}^{-1} \sum_{i=1}^{n} (X_i - \tilde{X}_i)(Y_i - \tilde{Y}_i) \mathbf{1}_i$$

- $n \to \infty$ のときに $b \to 0$ とすればよさそうである。
- (このような配慮は漸近理論のために必要なのであるが、実際に使う際には無視されることが多く、それによる影響も小さい。)

$\hat{\beta}$ の漸近正規性

- 「実行不可能な」推定量 $\hat{\beta}_{\mathrm{inf}}$ の漸近正規性は既に示した:

$$\sqrt{n}(\hat{\beta}_{\mathrm{inf}} - \beta) \xrightarrow{d} N(0, \Phi^{-1}\Psi\Phi^{-1})$$

- 追加的にいくつかの条件が満たされれば、「実行可能な」推定量 $\hat{\beta}$ も漸近正規性を持つ:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Phi^{-1}\Psi\Phi^{-1})$$

- 漸近分布が同じ!

Density-weighted approach

- トリミングを行うには,、パラメータ b を選ぶ必要がある (どちみち $b \to 0$ とするなら問題ナシ?)。
- ランダムな分母問題を解決するもう一つの方法として、Density-weighted approach がある。
- アイデア:
  分母 $f_i = f(Z_i)$ が問題ならあらかじめ払ってしまおう。

$$(Y_i - E[Y_i|Z_i])f_i = (X_i - E[X_i|Z_i])'\beta f_i + u_i f_i$$

後は OLS により β が推定できる (推定量を $\hat{\beta}_{\inf,f}$ とする)。

$\hat{\beta}_{\text{inf},f}$ の漸近正規性

- $\hat{\beta}_{\text{inf},f}$ は漸近正規性を持つ:

$$\sqrt{n}(\hat{\beta}_{\text{inf},f} - \beta) \xrightarrow{d} N(0, \Phi_f^{-1}\Psi_f\Phi_f^{-1})$$

$$\Phi_f = E[\tilde{X}_i\tilde{X}_i' f_i^2]$$

$$\Psi_f = E[\sigma^2(X_i, Z_i)\tilde{X}_i\tilde{X}_i' f_i^4]$$

$\hat{\beta}_f$ の漸近正規性

- $E[Y_i|Z_i], E[X_i|Z_i], f_i$ の (ノンパラ) 推定量をプラグインした上での推定量を $\hat{\beta}_f$ とする。
- いくつかの条件を加えると、$\hat{\beta}_f$ も漸近正規性を持つ:

$$\sqrt{n}(\hat{\beta}_f - \beta) \xrightarrow{d} N(0, \Phi_f^{-1}\Psi_f\Phi_f^{-1})$$

- 漸近分布が同じ!

ノンパラパートの推定1

- ここまで、パラメトリックパート (β) の推定を行ってきた。そこで得られた推定量 $\hat{\beta}$ を用いてノンパラパート $g(z)$ の推定を行う。
- モデルの式から $g(Z_i) = E[Y_i - X_i'\beta|Z_i]$ なので、以下のような $g(z)$ の推定量が考えられる。

$$\tilde{g}(z) = \frac{\sum_{j=1}^{n}(Y_j - X_j'\beta)K_h(z, Z_j)}{\sum_{j=1}^{n} K_h(z, Z_j)}$$

ノンパラパートの推定2

- β は未知なので、その推定量 $\hat{\beta}$ で代用した以下のような推定量が考えられる。

$$\hat{g}(z) = \frac{\sum_{j=1}^{n}(Y_j - X_j'\hat{\beta})K_h(z, Z_j)}{\sum_{j=1}^{n} K_h(z, Z_j)}$$

- $\hat{\beta}$ が $\sqrt{n}$ 一致性を持つのであれば、$\hat{\beta}$ の収束レートは他のどんなノンパラ推定量よりも速い。従って β を $\hat{\beta}$ で置き換えてもノンパラパートの推定量の漸近的性質に影響を与えない。

パラメトリックパートとノンパラパートの一致性

- $\hat{\beta}$ が $\sqrt{n}$ 一致性を持つには、$E[Y_i|Z_i], E[X_i|Z_i]$ などの推定に高次カーネルを用いる必要がある。
- $\hat{\beta}$ が $\sqrt{n}$ 一致性を持つのであれば、$g(z)$ の推定に高次カーネルを用いる必要はない (2 次カーネルが良い?)。

バンド幅hの選択

- ノンパラメトリックモデルのときと同様に、バンド幅hは最小二乗クロスバリデーションによって選ぶとよい。
- 具体的には、

$$\sum_{i=1}^{n}\left[Y_i - X_i'\hat{\beta} - \hat{g}_{-i}(Z_i, h)\right]$$

が最小になるように$h_1, ..., h_q$を選択する。

Andrews の MINPIN

- Andrews(1994) はセミパラメトリック推定量の $\sqrt{n}$ 一致性と漸近正規性に関する一般的なフレームワークを与えている。
- **P**reliminary **I**nfinite dimensional **N**uisance parameter に依存しうる目的関数 **MIN**imize することから **MINPIN** と呼ばれる。

MINPINの手順

- $\theta \in \mathbb{R}^p$ を関心のある有限次元パラメータ、$\tau \in \mathcal{T}$($\mathcal{T}$ はなめらかな関数の集合) を無限次元の局外パラメータとする。
- まず、局外パラメータ τ を (ノンパラメトリック) 推定する。
- 次に、τ の推定量を含んだ目的関数を最小にするように、興味のある θ を推定する。

A criterion function that may depend on Preliminary Infinite dimensional Nuisance parameter

- セミパラメトリックモデルは、確率変数を W として真のパラメータ (θ_0, τ_0) の下で1階の条件から得られるモーメント条件

$$E[m(W_i, \beta_0, \tau_0)] = 0$$

  を満たすとする。
- ここで m は θ に関して微分可能とする。

A criterion function that may depend on Preliminary Infinite dimensional Nuisance parameter estimator

- τ_0 のノンパラ一致推定量を $\hat{\tau}$ とする。
- θ_0 の推定量 $\hat{\theta}$ も、$\hat{\tau}$ に依存する最小化の 1 階の条件

$$\sqrt{n}\bar{m}_n(\hat{\theta}, \hat{\tau}) = 0$$

$$\bar{m}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} m(W_i, \theta, \hat{\tau})$$

を満たすとする。

漸近理論へ

- 1 階の条件
$$\sqrt{n}\bar{m}_n(\hat{\theta}, \hat{\tau}) = 0$$
を θ に関して θ_0 まわりでテイラー展開すると
$$\sqrt{n}\bar{m}_n(\theta_0, \hat{\tau}) + \frac{\partial}{\partial \theta'}\sqrt{n}\bar{m}_n(\bar{\theta}, \hat{\tau})(\hat{\theta} - \theta_0) = 0$$
ただし、$\bar{\theta}$ は $\hat{\theta}$ と θ_0 の間の値。

漸近正規性 1

- いくつかの条件下で

$$
\begin{aligned}
\frac{\partial}{\partial \theta'} \bar{m}_n(\bar{\theta}, \hat{\tau}) &= \frac{1}{n} \sum_i \frac{\partial}{\partial \theta'} m(W_i, \bar{\theta}, \hat{\tau}) \\
&\xrightarrow{p} E\left[\frac{1}{n} \sum_i \frac{\partial}{\partial \theta'} m(W_i, \theta_0, \tau_0) \right] \\
&= E\left[\frac{\partial}{\partial \theta'} m(W_i, \theta_0, \tau_0) \right] \\
&= M
\end{aligned}
$$

漸近正規性 2

- M が正則行列ならば、

$$
\begin{aligned}
\sqrt{n}(\hat{\theta}-\theta_0) &= -\left(M^{-1}+o_p(1)\right)\frac{1}{\sqrt{n}}\sum_i m(W_i,\theta_0,\hat{\tau}) \\
&= -\left(M^{-1}+o_p(1)\right)\left\{\frac{1}{\sqrt{n}}\sum_i m(W_i,\theta_0,\tau_0)+\ldots\right\}
\end{aligned}
$$

- … は

$$
\sqrt{n}\bar{m}_n(\theta_0,\hat{\tau})-\sqrt{n}\bar{m}_n(\theta_0,\tau_0)
$$

漸近正規性 3

- ... はひとまず無視

$$
\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta_0) &= -(M^{-1} + o_p(1))\frac{1}{\sqrt{n}}\sum_i m(W_i, \theta_0, \tau_0) \\
&\xrightarrow{d} N(M^{-1}SM'^{-1})
\end{aligned}
$$

ただし

$$
S = V[m(W_i, \theta_0, \tau_0)]
$$

- ... が無視できる条件は?

決め手はStochastic Equicontinuity!

- 各 $\tau \in \mathcal{T}$ について経験過程 $\nu_n(\tau)$ を以下のように定義する。

$$\nu_n(\tau) = \sqrt{n}(\bar{m}_n(\theta_0, \tau) - E[\bar{m}_n(\theta_0, \tau)])$$

確率的同程度連続性 (Stochastic Equicontinuity)

任意の $\epsilon > 0$ と $\eta > 0$ に対して、$\delta > 0$ が存在して

$$\limsup_{n \to \infty} P^* \left[\sup_{\tau \in \mathcal{T}, \rho_{\mathcal{T}}(\tau, \tau_0) < \delta} |\nu_n(\tau) - \nu_n(\tau_0)| > \eta \right] < \epsilon$$

が成り立つとき、$\{\nu_n(\tau) : n \geq 1\}$ は τ_0 において**確率的同程度連続 (Stochastic Equicontinuous)** であるという

Stochastic Equicontinuity エラい 1

- $\hat{\tau}$ が τ_0 の一致推定量であり、$\nu_n(\tau)$ が τ_0 で Stochastic Equicontinuous であるという条件下で

$$
\begin{aligned}
&\limsup_{n\to\infty} P(|\nu_n(\hat{\tau}) - \nu_n(\tau_0)| > \eta) \\
\leq &\limsup_{n\to\infty} P(|\nu_n(\hat{\tau}) - \nu_n(\tau_0)| > \eta, \hat{\tau} \in \mathcal{T}, \rho_{\mathcal{T}}(\hat{\tau}, \tau_0) \leq \delta) \\
&+ \limsup_{n\to\infty} P(\hat{\tau} \notin \mathcal{T} \text{ or } \rho_{\mathcal{T}}(\hat{\tau}, \tau_0) > \delta) \\
\leq &\limsup_{n\to\infty} P^* \left[\sup_{\tau\in\mathcal{T}, \rho_{\mathcal{T}}(\tau,\tau_0)<\delta} |\nu_n(\tau) - \nu_n(\tau_0)| > \eta \right] < \epsilon
\end{aligned}
$$

- 従って $\nu_n(\hat{\tau}) - \nu_n(\tau_0) \xrightarrow{p} 0$

Stochastic Equicontinuityエラい2

- 無視した部分 (...) は

$$\nu_n(\hat{\tau}) - \nu_n(\tau_0) + \sqrt{n}E[\bar{m}_n(\theta_0, \hat{\tau})]$$

- 多くのセミパラモデルでは以下が成立 (漸近的な直交条件)

$$\sqrt{n}E[\bar{m}_n(\theta_0, \hat{\tau})] = o_p(1)$$

- 従って Stochastic Equicontinuity の下で

$$... = o_p(1)$$

Stichastic Equicontinuity エラい 3

- Stochastic Equicontinuity の下で

$$\sqrt{n}(\hat{\theta} - \theta_0) = (M^{-1} + o_p(1)) \left[\frac{1}{\sqrt{n}} \sum_i m(W_i, \theta_0, \tau_0) + o_p(1) \right]$$

なので

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, M^{-1} S M'^{-1})$$

1. 部分線形モデル
2. MINPIN 推定量
3. セミパラメトリック効率性限界
4. Appendix: 関数解析の基礎

パラメトリック効率性限界

- 有限次元の未知パラメータ θ の値以外には分布がわかっているパラメトリックモデルを考える。
- θ の推定量 $\hat{\theta}$ の漸近分散の下限は Fisher 情報行列 I の逆行列として与えられる。

$$I_{i,j} = E\left[\frac{\partial \log f(X,\theta)}{\partial \theta_i}\frac{\partial \log f(X,\theta)}{\partial \theta_j}\right]$$

セミパラをパラメトリックとして扱う

- ある分布 $P \in \mathcal{P}$ から無作為標本を得たとする。
- 分布 P により決まる k 次元のパラメータ $\theta = \psi(P)$ に関心がある。(ひとまず $k = 1$ の場合を考える。)
- 分布 P を含み、$\mathcal{P}$ に含まれる、パラメトリックモデルの集合

  $$\mathcal{P}_0 = \{P_\theta : \theta \in \Theta\} \subset \mathcal{P}$$

  に注目する。ここでの $\mathcal{P}_0$ をパラメトリックサブモデルという。
- パラメトリックサブモデルはパラメトリックモデルなので最尤推定が可能。

パラメトリック効率性限界とセミパラメトリック効率性限界の関係

- $\mathcal{P}_0 \subset \mathcal{P}$ なので、サブモデル全体 ($\mathcal{P}_0$) の中で $\theta = \psi(P)$ の推定を行うより、モデル全体 ($\mathcal{P}$) の中で推定を行う方が大変。
- 従って、セミパラメトリックモデル推定量の分散は少なくともパラメトリックサブモデルの推定量の分散以上となるはずである。
- よって、セミパラメトリック効率性限界を、すべてのパラメトリックサブモデルの中での分散の最大値として定義する。

パラメトリックサブモデルのスコア関数

- $P_t, t \in [0, \epsilon)$ をパラメトリックサブモデルとする。
- 真の分布 P を P_0 として、$g \in L_2(P)$ が存在して

$$\int \left[\frac{dP_t^{\frac{1}{2}} - dP^{\frac{1}{2}}}{t} - \frac{1}{2} g dP^{\frac{1}{2}} \right]^2 \to 0$$

を満たすとき、この g を P_t の**スコア関数**という。

- 分布 P_t が密度関数 p_t を持つなら

$$g(x) = \frac{\partial}{\partial t} \log p_t(x) |_{t=0}$$

接集合

- パラメトリックサブモデルごとにスコア関数が登場する。
- 真の分布 P を含むパラメトリックサブモデルのスコア関数の集合を、$\mathcal{P}$ の P における**接集合**といい、$\dot{\mathcal{P}}_P$ と表す。
- 特に、接集合が線形空間である場合、**接空間**と呼ばれる。

ψ の微分可能性

- $\psi : \mathcal{P} \to R^k$ が $P(= P_0)$ で (Gateaux) 微分可能とは、任意の $g \in \dot{\mathcal{P}}_P$ とそれをスコアとするパラメトリックサブモデル $P_t, t \in [0, \epsilon)$ について

$$\frac{\psi(P_t) - \psi(P)}{t} \to \dot{\psi}_P g$$

  を満たす連続な線形写像 $\dot{\psi}_P$ が存在する。
- 各パラメトリックサブモデルで ψ が P で微分可能であるとする。

接空間における内積

- 接集合 $\dot{\mathcal{P}}_P$ が線形空間 (つまり接空間) だとする。
- $\dot{\mathcal{P}}_P$ における内積を

$$< f, g >_P = \int f\bar{g}dP$$

として定義すると、$\dot{\mathcal{P}}_P$ は Hilbert 空間。

$\dot{\psi}_P$を内積により表現する

- 作用素 $\dot{\psi}_P$ は連続なので有界。
- Riesz の表現定理より、$\tilde{\psi}_P \in \dot{\mathcal{P}}_P$ が存在して

$$\dot{\psi}_P g = < \tilde{\psi}_P, g >_P = \int \tilde{\psi}_P g dP$$

が成り立つ。

唯一のパラメータ t を最尤推定する

- パラメトリックサブモデル $P_t, t \in [0, \epsilon)$ は1つのパラメータ t を持つパラメトリックモデルと見ることができる。
- パラメータ t は最尤法により推定することができ、その推定量を $\hat{t}$ とする。$\sqrt{n}\hat{t}$ の漸近分散は $E[g^2]^{-1}$ で与えられる。
- $\hat{t}$ を用いると、θ の最尤推定量は $\hat{\theta} = \psi(P_{\hat{t}})$ と表される。

$\psi(P_t)$ を t に関して展開する

- $\psi(P_t)$ を t に関して真の値 $t=0$ の周りで展開すると

$$\hat{\theta} = \psi(P_0) + (\dot{\psi}_P g)\hat{t} + \dots$$

- $\psi(P_0) = \theta$ に注意すると、

$$\sqrt{n}(\hat{\theta} - \theta) =< \tilde{\psi}_P, g >_P \sqrt{n}\hat{t} + o_p(1)$$

$\hat{\theta}$ の漸近分散の上限

- $\sqrt{n}\hat{t}$ の漸近分散は $< g, g >_P^{-1}$ なので、$\hat{\theta}$ の漸近分散は

$$\frac{< \tilde{\psi}_P, g >_P^2}{< g, g >_P}$$

これに Cauchy-Schwartz の不等式を適用すると

$$\frac{< \tilde{\psi}_P, g >_P^2}{< g, g >_P} \leq \frac{< \tilde{\psi}_P, \tilde{\psi}_P >_P < g, g >_P}{< g, g >_P} =< \tilde{\psi}_P, \tilde{\psi}_P >_P$$

が成り立つ。

セミパラメトリック推定量の効率性限界

- $\tilde{\psi}_P \in \dot{\mathcal{P}}_P$ より、g を $\dot{\mathcal{P}}_P$ の中で動かしたときの「スコア関数が g であるようなサブモデルの漸近分散」の上限が $< \tilde{\psi}_P, \tilde{\psi}_P >_P$ であることがわかる。
- 従って $< \tilde{\psi}_P, \tilde{\psi}_P >_P$ がセミパラメトリック推定量の効率性限界。
- パラメトリックサブモデルの分散の上限を達成するものを**最も不都合な**パラメトリックサブモデルという。

興味のあるパラメータ θ が多次元の場合

- $\psi(P) \in \mathbb{R}^k$ の場合、ベクトル $h \in \mathbb{R}^k$ をかけた $h^T\theta$ を考える。

$$\sqrt{n}(h^T\hat{\theta} - h^T\theta) = (h^T\dot{\psi}_P g)\sqrt{n}\hat{t} + o_p(1)$$

- $h^T\theta$ の漸近分散は

$$\frac{< h^T\tilde{\psi}_P, g >_P^2}{< g, g >_P} \leq \frac{< h^T\tilde{\psi}_P, h^T\tilde{\psi}_P >_P < g, g >_P}{< g, g >_P} =< h^T\tilde{\psi}_P, h^T\tilde{\psi}_P >_P$$

効率的影響関数

- 従って、θ が多次元の場合にも

$$\sqrt{n}(\hat{\theta} - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{\psi}_P(X_i) + o_p(1)$$

  となる(セミパラ)推定量 $\hat{\theta}$ が、分散の下限 $E[\tilde{\psi}_P \tilde{\psi}_P^T]$ を達成するため、最も効率的であると言える。
- $\tilde{\psi}_P$ は**効率的影響関数**と呼ばれる。

ノルム空間

ノルム

線形空間 X の任意のベクトル x に実数 $||x||$ が対応していて、以下の条件を満たすとき、X にノルムが定義されているといい、$||x||$ を x の**ノルム**という。

(N1) $||x|| \geq 0$ かつ、$||x|| = 0 \Leftrightarrow x = 0$
(N2) $||\alpha u|| = |\alpha|||x||$
(N3) $||x + y|| \leq ||x|| + ||y||$

ノルム空間

ノルムが定義されている線形空間を**ノルム空間**という。
ノルムにより距離を定義すればノルム空間は距離空間。

内積空間

内積

線形空間 X の任意のベクトル x, y, z と任意の $\alpha, \beta \in \mathbb{R}$ に対して、以下の条件を満足する $< x, y > \in \mathbb{R}$ が定まるとき、$< x, y >$ は x と y の**内積**という。
(H1)$< x, y > = \overline{< y, x >}$
(H2)$< x, x > \geq 0$ かつ、$< x, x > = 0 \Leftrightarrow x = 0$
(H3)$< \alpha x + \beta y, z > = \alpha < x, z > + \beta < y, z >$
内積を用いると $x \in X$ のノルムを $||x|| = \sqrt{< x, x >}$ と定義することができる。

内積空間

内積が定義されている線形空間を**内積空間**という。

Banach 空間・Hilbert 空間

完備距離空間

距離空間 (X, d) の任意のコーシー列に対して、その点列の極限が「X に属する点」として存在することが保証されるとき、X は**完備**であるという。

Banach 空間

完備なノルム空間を **Banach 空間**という。

Hilbert 空間

完備な内積空間 (内積によって誘導されるノルム空間) を **Hilbert 空間**という。

有界作用素と双対空間

有界作用素

線形写像 $T : X \to Y$ に対してある $r > 0$ が存在して

$$\forall x \in X, ||Tx|| \leq r||x||$$

が成立するとき、T は**有界作用素**という。
X から Y への有界作用素全体を $\mathcal{B}(X, Y)$ と表す。

双対空間

Y が Banach 空間の場合、$\mathcal{B}(X, Y)$ も Banach 空間である。
Banach 空間 $\mathcal{B}(X, \mathbb{R})$ を X の**双対空間**といい、X^* と表す。

Riesz の表現定理

Riesz の表現定理

$\mathcal{H}$ を Hilbert 空間とする。
任意の $\psi \in \mathcal{H}^*$ について、$x_\psi \in \mathcal{H}$ が唯一つ存在して、

$$\forall x \in \mathcal{H}, \psi(x) =< x, x_\psi >$$

さらに $||\psi|| = ||x_\psi||$ が成り立つ。