6. Bayesian Inverse Problem

竹田航太

2022 年 2 月 11 日

概要

観測値から真の状態を推定する「逆問題」を一般的に定式化する．Banach 空間の状態空間と観測空間を考え，その上の確率測度や作用素によってノイズや観測を表現する．関数解析や（Banach 空間上の）測度論，確率論の知識を必要とする．基本的に Sullivan[2] の Chapter6 に沿っているが Nickl[3] も大いに参考にした.

1 Inversion and Regularization

観測値の情報から真の状態を推定するような逆問題を考える．真の状態の空間 $\mathcal{H}$ と観測値の空間 $\mathcal{K}$ を考える．観測作用素 $A : \mathcal{H} \to \mathcal{K}$ とし，観測ノイズを確率変数 $\eta \in \mathcal{K}$ で表すと，$u \in \mathcal{H}$ に対して観測は $y = Au + \eta$ と表される．観測値 $y \in \mathcal{K}$ に対して推定値 $\hat{u} \in \mathcal{H}$ を与える線形作用素を $K : \mathcal{K} \to \mathcal{H}$ を考える．以上のような線形推定について「最適」な作用素を与える結果が以下の定理.

Definition 1.1 (不偏推定)**.** 観測モデル $y = Au + \eta$ に対する推定線形作用素 $K : \mathcal{K} \to \mathcal{H}$ が $\mathbb{E}[Ky] = u$ を満たすとき不偏推定と言う.

Definition 1.2 (推定誤差)**.** 推定値 $\hat{u}$ に対して以下を定義する.

- 平均二乗誤差 *(MSE):* $\mathbb{E}[\|\hat{u} - u\|_{\mathcal{H}}^2]$
- 誤差共分散作用素 *(Cov.op.):* $\mathbb{E}[(\hat{u} - u) \otimes (\hat{u} - u)]$

Definition 1.3 (作用素の大小)**.** 半正定値作用素 A, B に対して

$$A \geq B \overset{def}{\Leftrightarrow} A - B : \text{ 半正定値}$$

と定める.

Theorem 1.4 (Gauss Markov)**.** $\mathcal{H}, \mathcal{K}$ を可分 *Hilbert* 空間，$A : \mathcal{H} \to \mathcal{K}$ を線形作用素とする．状態変数 $u \in \mathcal{H}$ に対し，観測 $y = Au + \eta$ とするモデルを考える．ただし，$\eta \in \mathcal{K}$ は平均 0, 誤差共分散が対称正定値作用素 $Q : \mathcal{K} \to \mathcal{K}$ であるようなノイズとする．また，$A^*Q^{-1}A$ は可逆と

する．

このとき，不偏線形推定 $K : \mathcal{K} \to \mathcal{H}$ の中で *MSE* と *Cov.op.* を同時に最小化する意味で最適な作用素は

$$K = (A^*Q^{-1}A)^{-1}A^*Q^{-1}$$

である．さらに，$\hat{u} = Ky$ は以下を満たす．

(1) $\mathbb{E}[\hat{u}] = u$

(2) $\mathbb{E}[(\hat{u} - u) \otimes (\hat{u} - u)] = (A^*Q^{-1}A)^{-1}$

Remark 1.5. *Theorem1.4* で有限次元の場合に $A^*Q^{-1}A$ は可逆でない場合には逆行列を擬似逆行列 *(Moore-Penrose* の *pseudo inverse)* で代用できる．

$$K = (A^*Q^{-1}A)^{\dagger}A^*Q^{-1}$$

有限次元で Q の次数が退化している場合にはこの擬似逆行列や正則化が有効である．

2 Bayesian Inversion

Banach 空間上の逆問題に対して，ベイズを用いた定式化を行う．

Theorem 2.1 (generalized bayes rule)**.** $\mathcal{U}$*: Banach* 空間，$H : \mathcal{U} \to \mathbb{R}^m$ 連続，η*: r.v. on* $\mathbb{R}^m$ はルベーグ測度に対する密度関数 ρ をもつ確率測度に従うとする．u*: r.v. on* $\mathcal{U}$ は μ_0 に従い，$y = H(u) + \eta$ とする．

このとき，$P(u|y)$ は次を満たす μ^y で与えられる．$\mu^y \ll \mu_0$ かつ

$$\frac{d\mu^y}{d\mu_0}(u) \propto \exp(-\Phi(u; y))$$

ただし，$\Phi(u; y) = -\log(\rho(y - H(u))) + f(y)$，$f(y)$ は任意の加法的関数.

Remark 2.2. 定理の解釈について

- 状態変数 $u \in \mathcal{U}$ から観測関数 H とノイズ η によって得られる観測値 y から u を復元する際のベイズの定理を表している．
- $\rho \to \Phi$ の決め方の自由度から μ^y は正規化されていないことがわかる．
- ノイズ η は正規分布とは限らない．

3 wellposedness and Approximation

$\mathcal{U}, \mathcal{Y}$ を Banach 空間とし，$u \sim \mu_0$ on $\mathcal{U}$ とする．ここでは正規分布の事前分布を考えて（i.e. μ_0 は正規分布），前節とは異なり与えられたポテンシャル Φ から事後分布を定義することを考える．

Assumption 3.1 (Φ への制約)**.** *Φ に次の制約を定める．*

A1) $\forall \epsilon, r > 0, \exists M = M(\epsilon, r) \in \mathbb{R}$ *s.t. for all* $u \in \mathcal{U}, \|y\|_{\mathcal{Y}} < r$
$\Phi(u; y) \geq M - \epsilon \|u\|_{\mathcal{U}}$

A2) $\forall r > 0, \exists K = K(r) > 0$ *s.t. for all* $\|u\|_{\mathcal{U}}, \|y\|_{\mathcal{Y}} < r$
$\Phi(u; y) \leq K(r)$

A3) $\forall r > 0, \exists L = L(r) > 0$ *s.t. for all* $\|u_1\|_{\mathcal{U}}, \|u_2\|_{\mathcal{U}}, \|y\|_{\mathcal{Y}}$
$|\Phi(u_1; y) - \Phi(u_2; y)| \leq L \|u_1 - u_2\|_{\mathcal{U}}$

A4) $\forall \epsilon, r > 0, \exists C = C(\epsilon, r) > 0$ *s.t. for all* $\|u\|_{\mathcal{U}}, \|y_1\|_{\mathcal{Y}}, \|y_2\|_{\mathcal{Y}}$
$|\Phi(u; y_1) - \Phi(u; y_2)| \leq \exp(\epsilon \|u\|_{\mathcal{U}}^2 + C) \|y_1 - y_2\|_{\mathcal{Y}}$

Remark 3.2. *観測関数 H に 2 つの条件をつけることで前節において H から定義される Φ はこれらの制約を満たすことが示される [1].*

次の定理で Φ から定まる事後分布が well-defined であることが保証される．

Theorem 3.3 (welldefined posterior)**.** *Φ は A1)　A3) を満たし，μ_0 は Banach 空間 $\mathcal{U}$ 上の正規分布とする．*

このとき，任意の $y \in \mathcal{Y}$ に対して μ^y を次で定義すると well-defined な $\mathcal{U}$ 上の確率測度となる．

$$\frac{d\mu^y}{d\mu_0}(u) = Z(y)^{-1} \exp(-\Phi(u; y))$$

ただし，$Z(y) = \int_{\mathcal{U}} \exp(-\Phi(u; y)) d\mu_0(u)$

仮定 A3) がないと μ^y の有界性が示せないが，差の bound は与えることができる．

Theorem 3.4 (locally Lipschitz)**.** *$\Phi(u; y)$ は A1),A2),A4) を満たし，μ_0 は $\mathcal{U}$ 上の正規分布とする．任意の $y \in \mathcal{Y}$ に対し，$\mu^y \ll \mu_0$ かつ $Z(y)$ を正規化定数として*

$$\frac{d\mu^y}{d\mu_0}(u) = Z(y)^{-1} \exp(-\Phi(u; y))$$

が成り立つとする．

このとき，$\forall r > 0$ に対して $\exists C(r) > 0$ $s.t.$ $\forall y, y' \in B_r(0)$

$$d_H(\mu^y, \mu^{y'}) \leq C(r)\|y - y'\|_{\mathcal{Y}}$$

が成り立つ．ただし，$B_r(0) = \{y \in \mathcal{Y}; \|y\|_{\mathcal{Y}} \leq r\}$ であり，$d_H(\cdot,\cdot)$ は *Hellinger* 距離.

次に離散モデル化による誤差のベイズ推定への影響を評価する．離散化パラメータ $N \in \mathbb{N}$ を導入する.

Assumption 3.5 (離散モデル化誤差)**.** *A5)* $\forall \epsilon > 0, \exists K(\epsilon)$ $s.t.$ $\forall u \in \mathcal{U}, \forall y \in \mathcal{Y}$

$$|\Phi(u;y) - \Phi^N(u;y)| \leq K \exp(\epsilon \|u\|_{\mathcal{U}}^2)\psi(N)$$

ただし，$\psi(N) \to 0(\mathbb{N} \to \infty)$ が成り立つとする.

Theorem 3.6 (locally Lipschitz2)**.** 離散化パラメータ $N \in \mathbb{N}$ に対して, Φ, Φ^N は *A1),A2),A5)* を満たすとする（N によらず一様）．さらに，$\mu_N^y \ll \mu_0$ を正規化定数 $Z(y)^N$ を使って

$$\frac{d\mu_N^y}{d\mu_0}(u) = (Z(y)^N)^{-1}\exp(-\Phi^N(u;y))$$

で定める.

このとき, 任意の $r > 0$ に対して, N に依存しない $C(r) > 0$ が存在して次を満たす．$\forall \|y\| < r$ で

$$d_H(\mu^y, \mu_N^y) \leq C(r)\psi(N)$$

4 Frequentist Consistency of Bayesian Method

Bayes 推定においては事後分布が事前分布に強く依存し（特に Banach 空間上では)，データの情報が反映されづらい．Bayes 推定の結果に対してその「正しさ」を頻度論の立場から考えてみる．この節では estimator の列の収束を通して「正しさ」を議論する.

Definition 4.1. $\mathcal{U}, \mathcal{Y}$*: Banach* 空間*,* $\mathcal{M}_1(\mathcal{Y})$: $\mathcal{Y}$ 上の確率測度の全体*,* 観測データ $\{Y_1, \ldots, Y_n\} \subset \mathcal{Y}$ が与えられているとする.

(1) $\mu^\dagger \in \mathcal{M}_1(\mathcal{Y})$ が観測データ $\{Y_1, \ldots, Y_n\} \subset \mathcal{Y}$ の真の分布
$\overset{def}{\Leftrightarrow}$ $Y_1, \ldots, Y_n$ は分布が $\mu^\dagger$ の *iid*
(2) $u \in \mathcal{U}$ が与えられたときの y の分布を $\mu(\cdot|u) \in \mathcal{M}_1(\mathcal{Y})$ として確率測度の族 $\{\mu(\cdot|u) \in \mathcal{M}_1(\mathcal{Y}); u \in \mathcal{U}\}$ を尤度モデルと呼ぶ.
密度関数が存在する場合はその族を尤度モデルと呼ぶこともある.

(3) 尤度モデルが *well-specified* であるとは以下が成り立つことを言う．
$\exists u^\dagger \in \mathcal{U}\ s.t.\ \mu^\dagger = \mu(\cdot|u^\dagger)$
また，成り立たないとき *misspecified* という．

Assumption 4.2. 以下を仮定する．

(1) $\mathcal{U} \subset \mathbb{R}^p, \mathcal{Y} \subset \mathbb{R}^q$ とする
(2) 各 $u \in \mathcal{U}$ で $\mu(\cdot|u)$ はルベーグ速度に対して絶対連続である．
(3) 尤度モデルは *well-specified* であるとする．
(4) ある $f(\cdot|\cdot) : \mathcal{Y} \times \mathcal{U} \to [0, \infty)\ s.t.$
- 各 $u \in \mathcal{U}$ で $\mu(E|u) = \int_E f(y|u)dy\ \forall E : \mathcal{Y}$ の可測集合
- 各 $y \in \mathcal{Y}$ で $f(y|\cdot) : \mathcal{U} \to \mathbb{R}$ は連続

Definition 4.3. *(*観測データ $Y_1, \cdots Y_n$ を用いた *MLE)* 観測データ：$\{Y_1, \cdots, Y_n\} \subset \mathcal{Y}$ に対して，

(1) 対数尤度関数を次で定義する．
$L_n(u; Y_1, \cdots, Y_n) = -\frac{1}{n}\sum_{i=1}^n \log(f(Y_i|u))$
(2) 最尤推定 *(Maximum Likelihood Estimator: MLE)* $\hat{u}_n$ を次で定める．
$\hat{u}_n \in argmax_{u\in\mathcal{U}} \prod_{i=1}^n f(Y_i|u) = argmin_{u\in\mathcal{U}} L_n(u; Y_1, \cdots, Y_n)$
(3) 真の分布 $\mu^\dagger$ に対して $\log(f(\cdot|u))$ が可積分であるとき，連続的な対数尤度関数 L を次で定める．
$L(u) = \mathbb{E}_{Y\sim\mu^\dagger}[-\log(f(Y|u))]$
(4) $\{f(\cdot|u); u \in \mathcal{U}\}$ が *identicable* とは次を満たすことを言う．
$f(\cdot|u_1) = f(\cdot|u_2)\ a.e. \Leftrightarrow u_1 = u_2$

4.1 Consistency

MLE の一貫性について確率変数の列の収束を通して考える．まず，仮想的な連続データ $Y \sim \mu^\dagger$ が得られた場合の一貫性について示す．

Lemma 4.4 (（連続的）MLE の一貫性)**.** *f が identicable であるとき，$u^\dagger \in \mathcal{U}$ に対して，$\mu^\dagger = \mu(\cdot|u^\dagger)$ ならば $u^\dagger$ は L の一意的な最小点となる．*

Theorem 4.5 (Consistency of MLE)**.** $f(y|u) > 0 \forall(u, y) \in \mathcal{U}\times\mathcal{Y}, \mathcal{U}$: コンパクト, $\{f(\cdot|u); u \in \mathcal{U}\}$: *identificable*, $\int_{\mathcal{Y}} \sup_{u\in\mathcal{U}} |\log(f(y|u))| f(y|u^\dagger)dy < \infty$ とする．
このとき，

$$\hat{u}_n \xrightarrow{P} u^\dagger\ (n \to \infty)$$

i.e. $\forall \epsilon > 0, \mathbb{P}[|\hat{u}_n - u^\dagger| \geq \epsilon] \to 0\ (n \to \infty)$

証明には次の補題を使う．

Lemma 4.6 (Theorem 1 in Nickl[3])**.** *$\mathcal{U} \subset \mathbb{R}^p$: コンパクト，$L: \mathcal{U} \to \mathbb{R}$: 連続（上で定義した L と同じとは限らない．），$u^\dagger$: L の unique minimizer とし，$\sup_{u\in\mathcal{U}} |L_n(u; Y_1, \cdots, Y_n) - L(u)| \xrightarrow{P} 0\ (n \to \infty)$ とする．*

このとき，$\forall \hat{u}_n$: MLE from $\{Y_1, \cdots, Y_n\}$ に対し，次が成り立つ．

$$\hat{u}_n \xrightarrow{P} u\ (n \to \infty)$$

4.2 Locally Asymptotic Behavior

Definition 4.7 (Fisher 情報行列)**.** $\{f(\cdot|u); u \in \mathcal{U}\}$: 尤度モデルとする．$u \in \mathcal{U} \subset \mathbb{R}^p$ に対して，*Fisher* 情報行列 $i_F(u) \in \mathbb{R}^p \times \mathbb{R}^p$ を以下で定める．

$$i_F(u) \coloneqq \mathbb{E}_{Y\sim f(\cdot|u)}[\nabla_u \log f(Y|u) \nabla_u \log f(Y|u)^\top] \tag{4.1}$$

Assumption 4.8 (Regularity Assumption)**.** $f: \mathcal{Y} \times \mathcal{U} \to [0, \infty)$

(a) $\forall u \in \mathcal{U}, \forall y \in \mathcal{Y}, f(y|u) > 0$
(b) $\exists u^\dagger \in int(\mathcal{U})$ *s.t.* $\mu^\dagger = \mu(\cdot|u^\dagger)$
(c) $\exists U \subset \mathcal{U};$ 開集合 *s.t.* $\forall y \in \mathcal{Y}, f(y|\cdot) \in C^2(U)$
(d) $\mathbb{E}[\nabla_u^2 \log f(Y|u^\dagger)|_{u=u^\dagger}] \in \mathbb{R}^p \times \mathbb{R}^p$ は正則行列で
$\mathbb{E}[\|\nabla_u \log f(Y|u)|_{u=u^\dagger}\|^2] < \infty$
(e) $\exists r > 0$ *s.t.* $B = B_r(u^\dagger) \subset U$ 以下が成り立つ

$$\mathbb{E}[\sup_{u\in B} \|\nabla_u^2 \log f(Y|u)\|] < \infty$$
$$\int_{\mathcal{Y}} \sup_{u\in B} \|\nabla_u \log f(Y|u)\| dy < \infty$$
$$\int_{\mathcal{Y}} \sup_{u\in B} \|\nabla_u^2 \log f(Y|u)\| dy < \infty$$

ただし，期待値については $\mathbb{E}[g(Y)] = \mathbb{E}_{Y\sim\mu^\dagger}[g(Y)] = \int_{\mathcal{Y}} g(y) f(y|u) dy$ ($\forall g$: 可測関数) とする．ノルムはベクトル，行列に対して各要素の *2* 乗和の平方根とする．

Theorem 4.9 (Local asymptotic normality of MLE)**.** *f は Regularity Assumption4.8 を満たすとする．また，$\hat{u}_n \xrightarrow{P} u\ (n \to \infty)$ とする．*

このとき以下が成り立つ．

(1) $i_F(u^\dagger) = \mathbb{E}_{Y\sim\mu^\dagger}[\frac{\partial^2 \log(f(Y|u))}{\partial u_i \partial u_j}|_{u=u^\dagger}]$
(2) $\sqrt{n}(\hat{u}_n \to u^\dagger) \xrightarrow{d} N(0, i_F(u^\dagger)^{-1})\ (n\to\infty)$

証明は Nickl[3] を参照.

Theorem 4.10 (Berstein von Mise)**.** *f は Regularity Assumption4.8 を満たし, $\mu_0 \in \mathcal{M}_1(\mathcal{U})$ はルベーグ速度に対して絶対連続で $u^\dagger \in supp(\mu_0)$ とする. さらに, 以下のような uniform consistency estimator T_n の存在を仮定する. $T_n : \mathcal{Y}^n \to \mathbb{R}^p$ s.t. $\forall\epsilon > 0$*

$$\sup_{u\in\mathcal{U}} \mathbb{P}_{Y_i\sim f(\cdot|u)}(\|T_n(Y_1,\cdots,Y_n) - u\| > \epsilon) \to 0\ (n\to\infty)$$

また, u の事後分布を $\mu_n = \mu_0(\cdot|Y_1,\cdots,Y_n)$ とかく.
このとき, 次が成り立つ.

$$d_{TV}\left(\mu_n, N(\hat{u}_n, i_F(u^\dagger)^{-1}/n)\right) \xrightarrow{P} 0\ (n\to\infty)$$

Remark 4.11. *証明には Parametric Testing Theory や Local Asymptotic Normality, Contiguity といった概念が必要である [3]. また, Prohorov の定理や Markov の不等式などの確率論の定理を利用する.*

4.3 補足

4.3.1 Cramér-Rao

Theorem 4.12 (Cramér-Rao)**.** *f : Regularity Assumption4.8 を満たすとする. 観測データは $Y_1,\ldots,Y_n$ の $\mu(\cdot|u)$ に従い, $\hat{u}_n = \hat{u}_n(Y_1,\ldots,Y_n)$ を u の不遍推定とする. (i.e $\mathbb{E}_u[\hat{u}_n] = u$) このとき任意の $u \in int(U)$ に対して*

$$Var[\hat{u}_n] \geq \frac{1}{n} i_F(u)^{-1} \tag{4.2}$$

ただし, 行列に対する不等式は差が正定値行列であることを使って定義する.

Proof. $n \in \mathbb{N}$ を固定し, $F(Y|u) := \prod_{i=1}^n f(Y_i|u)$, $l^{'}(u,Y) := \frac{d}{du}\log F(Y|u) = \sum_{i=1}^n \frac{d}{du}\log f(Y_i|u)$ とおく. (e) の 2 つ目と DCT より任意の $u \in int(B)$ で

$$0 = \frac{d}{du}\int_{\mathcal{Y}} f(y|u)dy = \int_{\mathcal{Y}} \frac{d}{du} f(y|u)dy = \int_{\mathcal{Y}} \frac{d}{du}\log f(y|u) f(y|u)dy$$

したがって

$$\mathbb{E}[l^{'}(u,Y)] = 0 \tag{4.3}$$

まず, 次元を $p=1$ として, Cauchy-Schwarz の不等式から

$$\begin{aligned} Cov_u^2(\hat{u}_n, l^{'}(u,Y)) &\leq Var_u[\hat{u}_n] Var_u[l^{'}(u,Y)] \\ &= Var_u[\hat{u}_n] n \cdot i_F(u) \end{aligned}$$

最後に $Cov_u^2(\hat{u}_n, l^{'}(u, Y)) = 1$ を示せば良い．Cov を計算すると (4.3) から

$$
\begin{aligned}
Cov_u^2(\hat{u}_n, l^{'}(u, Y)) &= \int_{\mathcal{Y}^n} (\hat{u}_n - u)\frac{d}{du}\log F(y|u)F(y|u)dy \\
&= \int_{\mathcal{Y}^n} (\hat{u}_n - u)\frac{d}{du}F(y|u)dy \\
&= \int_{\mathcal{Y}^n} \hat{u}_n(y)\frac{d}{du}F(y|u)dy - \int_{\mathcal{Y}^n} u\frac{d}{du}F(y|u)dy \\
&= \frac{d}{du}\int_{\mathcal{Y}^n} \hat{u}_n(y)F(y|u)dy - \int_{\mathcal{Y}^n} u\frac{d}{du}F(y|u)dy
\end{aligned}
$$

$\hat{u}_n$は不偏推定より

$$
\begin{aligned}
&= \frac{d}{du}\int_{\mathcal{Y}^n} uF(y|u)dy - \int_{\mathcal{Y}^n} u\frac{d}{du}F(y|u)dy \\
&= \int_{\mathcal{Y}^n} F(y|u)dy + \int_{\mathcal{Y}^n} u\frac{d}{du}F(y|u)dy - \int_{\mathcal{Y}^n} u\frac{d}{du}F(y|u)dy \\
&= \int_{\mathcal{Y}^n} F(y|u)dy \\
&= 1
\end{aligned}
$$

次に一般の $p \in \mathbb{N}$ の場合 $\epsilon = \hat{u}_n - u, \lambda = \frac{d}{du}\log F(Y|u), V = Var_u(\hat{u}_n), I = i_F[u], E : p$ 次単位行列として以下を示せば良い

(i)

$$
E\left[\begin{bmatrix}\epsilon \\ \lambda\end{bmatrix}\begin{bmatrix}\epsilon^\top & \lambda^\top\end{bmatrix}\right] \geq 0
$$

(ii)

$$
LHS = \begin{bmatrix} V & E \\ E & I \end{bmatrix}
$$

(iii)

$$
= \begin{bmatrix} E & I^{-1} \\ O & E \end{bmatrix}\begin{bmatrix} V - I^{-1} & E \\ E & I \end{bmatrix}\begin{bmatrix} E & O \\ I^{-1} & E \end{bmatrix}
$$

(iv) A : 正定値，正則, $A^\top BA$: 正定値 $\Rightarrow B$: 正定値

(v)

$$
\begin{pmatrix} A & O \\ O & B \end{pmatrix} : 正定値
$$

$\Rightarrow A$: 正定値

(i) 自明

(ii)$\mathbb{E}_u[(\hat{u}_n - u)\nabla_u \log F(Y|u)^\top] = E$ を示す．LHS の (i, j) 成分は $p = 1$ での Cov の計算と

$\frac{\partial u_i}{\partial u_j} = \delta_{i,j}$ から

$$\mathbb{E}_u[(\hat{u}_n - u)_i \frac{\partial}{\partial u_j} \log F(Y|u)] = \int_{\mathcal{Y}} \frac{\partial u_i}{\partial u_j} F(y|u) dy = \delta_{i,j}$$

となるので示された.
(iii) 計算する

$$\begin{aligned} RHS &= \begin{bmatrix} E & I^{-1} \\ O & E \end{bmatrix} \begin{bmatrix} V - I^{-1} & E \\ E & I \end{bmatrix} \begin{bmatrix} E & O \\ I^{-1} & E \end{bmatrix} \\ &= \begin{bmatrix} V - I^{-1} & E \\ O & I \end{bmatrix} \begin{bmatrix} E & O \\ I^{-1} & E \end{bmatrix} \\ &= \begin{bmatrix} V & E \\ E & I \end{bmatrix} \end{aligned}$$

(iv) 背理法により示す．$\exists y$s.t. $y^\top By < 0$ とする．$x = A^{-1}y$ とおくと仮定より $x^\top(A^\top BA)x \geq 0$．しかし，背理法の仮定から $x^\top(A^\top BA)x = y^\top By < 0$ より矛盾.
(v)$\forall x \in \mathbb{R}^p$ に対して，$z = (x, 0) \in \mathbb{R}^p \times \mathbb{R}^p$ をとると

$$0 \leq z^\top \begin{pmatrix} A & O \\ O & B \end{pmatrix} z = x^\top Ax$$

□

4.3.2 Bernstein Von Mise への準備

Parametric Testing Theory と Local Asymptotic Normality

参考文献

[1] A. M. Stuart. Inverse problems: A bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[2] Timothy John Sullivan. *Introduction to uncertainty quantification*, volume 63. Springer, 2015.

[3] Richard Nickl. Statistical theory. *Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge*, 2013.