GAUSSIAN SEQUENTIAL

$$y_j = \theta_j + \varepsilon \xi_j \qquad j = 1, 2, \cdots$$

$$\hat{\theta}(\lambda) = (\hat{\theta}_1, \hat{\theta}_2, \cdots). \quad \hat{\theta}_j = \lambda_j y_j.$$

Linear estimator

$$\lambda = \{\lambda_j\}_{j=1}^{\infty} \in \ell^2(\mathbb{N})$$

$$R(\lambda, \theta) = \mathbb{E}_\theta \|\hat{\theta}(\lambda) - \theta\|^2 = \sum_{j=1}^{\infty} \left[(1-\lambda_j)^2 \theta_j^2 + \varepsilon^2 \lambda_j^2 \right]$$

$\underline{\lambda \in \Lambda \subseteq \ell^2(\mathbb{N})}$. WEIGHT CLASS.

$$\lambda^{oracle}(\Lambda, \theta) = \arg\min_{\lambda \in \Lambda} R(\lambda, \theta)$$

UNBIASED ESTIMATION OF THE RISK.

$$\|\hat{\theta}(\lambda) - \theta\|^2 = \sum_j \left(\lambda_j^2 y_j^2 - 2\lambda_j y_j \theta_j + \theta_j^2 \right)$$

$$J(\lambda) \triangleq \sum_j \left(\lambda_j^2 y_j^2 - 2\lambda_j (y_j^2 - \varepsilon^2) \right)$$

$$\cancel{R(\lambda,\theta)} = \mathbb{E}(J(\lambda)) = \cancel{\mathbb{E}\sum_j \left(\lambda_j^2 y_j^2 - 2\lambda_j (y_j^2 - \varepsilon^2)\right) = \mathbb{E}\sum_j \left[(\lambda_j^2 - 2\lambda)(\theta_j^2 + \varepsilon^2 \xi_j^2 + 2\varepsilon\theta_j \xi_j) \right.} + 2\lambda_j \varepsilon^2 \Big]$$

$$\cancel{= \sum_j 2\lambda_j \varepsilon^2 + \sum_j (\lambda_j^2 - 2\lambda_j)(1+\varepsilon^2).}$$

$$\cancel{= \sum_j \lambda_j \left((\lambda_j - 2)(1+\varepsilon^2) + \varepsilon^2 \right). - \sum_j \lambda_j \left(\lambda_j + \lambda_j \varepsilon^2 - 2 - \theta\varepsilon^2 \right)}.$$

$$\cancel{-2\lambda_j y_j \theta_j} \qquad -\lambda_j \cancel{(y_j + \varepsilon)(y_j - \varepsilon)}.$$

$$\cancel{-\lambda_j (y_j + \varepsilon)}$$

$$\cancel{-\lambda_j ((\xi+1)\varepsilon + \theta_j)((\xi - 1)\varepsilon + \theta_j)}.$$

$$\mathbb{E}_\theta \|\hat{\theta}(\lambda) - \theta\|^2 = R(\lambda, \theta) = \mathbb{E}_\theta [J(\lambda)] + \sum_j \theta_j^2$$

$$\underline{\tilde{\lambda} = \tilde{\lambda}(\Lambda) = \arg\min_{\lambda \in \Lambda} J(\lambda)}$$

Risk mini

$\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots)$ random sequence whose $\tilde{\lambda}_j = \tilde{\lambda}_j(y)$ depend on all $y = (y_1, y_2, \cdots)$.

esti. weights $\tilde{\theta}(\Lambda) = \hat{\theta}(\tilde{\lambda}) = \{\tilde{\theta}_j\}$ $\quad \tilde{\theta}_j = \lambda_j(y) y_j$.

$\tilde{\theta}$ is a nonlinear estimator

$$\tilde{\theta}(\Lambda) \longrightarrow \cancel{\theta^{oracle}}\ \hat{\theta}(\lambda^{oracle}(\Lambda, \theta)).$$

ADAPTIVE TO ORACLE

$\Theta \subseteq \ell^2(\mathbb{N})$. $\Lambda \subseteq \ell^2(\mathbb{N})$
sequence weights

θ_ε^* : $\exists\, C < \infty$. s.t.

$$\mathbb{E}_\theta \|\theta_\varepsilon^* - \theta\|^2 \le C \inf_{\lambda \in \Lambda} \mathbb{E}_\theta \|\hat{\theta}(\lambda) - \theta\|^2$$

for all $\theta \in \Theta$. and $0 < \varepsilon < 1$.
θ_ε^* is esti for model with noise size ε.

exact sense. to oracle $\lambda^{oracle}(\Lambda, \cdot)$:

if $$\mathbb{E}_\theta \|\theta_\varepsilon^* - \theta\|^2 \le (1 + o(1)) \inf_{\lambda \in \Lambda} \mathbb{E}_\theta \|\hat{\theta}(\lambda) - \theta\|^2$$
$\forall\, \theta \in \Theta$ as $\varepsilon \to 0$ uniformly in $\theta \in \Theta$.

DESIGN THE CLASS Λ

1. only consider $\lambda_j \in [0, 1]$.
   projection of $\lambda_j \notin [0,1]$ on $[0, 1]$ will reduce risk $R(\lambda, \theta)$ of a linear estimator.

2. Usually it is sufficient to set $\lambda_j = 0$. for $j > N_{max}$.
$$N_{max} = \lfloor \tfrac{1}{\varepsilon^2} \rfloor$$
and it's computationally feasible

$\theta \in \Theta(\beta, Q)$. $\underline{\beta > \frac{1}{2}}$. $Q > 0$
Θ cont.

$$R(\lambda, \theta) = \sum_{j=1}^{N_{max}} \left[(1-\lambda_j)^2 \theta_j^2 + \varepsilon^2 \lambda_j^2 \right] + r_0(\varepsilon)$$

$$r_0(\varepsilon) = \sum_{j > N_{max}} \left[(1-\lambda_j)^2 \theta_j^2 + \underbrace{\varepsilon^2 \lambda_j^2}_{\text{vanish.}} \right] \qquad \lambda \in \ell^2(\mathbb{N}) \text{ and } N_{max} \to \infty$$

$$\le \sum_{j > N_{max}} \theta_j^2 + o(\varepsilon^2). \qquad \text{as } \varepsilon \to 0.$$

$$\le N_{max}^{-2\beta} \underbrace{\sum_{j > N_{max}} (j-1)^{2\beta} \theta_j^2}_{\text{part of const } Q} + o(\varepsilon^2).$$

$$= o(N_{max}^{-2\beta} + \varepsilon^2) = o(\varepsilon^2) \quad \text{as } \varepsilon \to 0.$$

choose $N = \lfloor \frac{1}{\varepsilon^2} \rfloor$ $\beta > \frac{1}{2}$. $o(\varepsilon^2)$.

E

EX1 FINITE DIMENSIONAL

$$\begin{cases} y_j = \theta_j + \varepsilon \xi_j & j = 1, \dots, d. \\ \xi_j \overset{iid}{\sim} N(0,1). \end{cases}$$

$$\Lambda_{const} = \{ \lambda \mid \lambda_j \equiv t,\ j = 1, \dots, d.\ t \in [0,1] \}.$$

$$\hat{\theta}(t) = \cancel{t(y)}\ ty$$

$$\min_t \mathbb{E}_\theta \| \hat{\theta}(t) - \theta \|^2 = \min_t \sum_{j=1}^{d} \left[(1-t)^2 \theta_j^2 + \varepsilon^2 t^2 \right]$$

$$= \frac{d\varepsilon^2 \|\theta\|^2}{d\varepsilon^2 + \|\theta\|^2}$$

$$t = t^* := \frac{\|\theta\|^2}{d\varepsilon^2 + \|\theta\|^2} \quad \text{minimize it.}$$

$$\lambda^{oracle}(\Lambda_{const}, \theta) = (t^*, \dots, t^*).$$

For $\lambda \in \Lambda_{const}$.

$$J(\lambda) = \sum_{j=1}^{d} \left(t^2 y_j^2 - 2t(y_j^2 - \varepsilon^2) \right) = (t^2 - 2t)\|y\|^2 + 2td\varepsilon^2$$

$= \frac{1}{n} \sum_{i=1} (\xi_i$ [illegible]

$+ \xi_i (p(X_i) - f(X_i))$

ASYMPTOTICALLY EFFICIENT

$$\tilde{t} = \left(1 - \frac{d\varepsilon^2}{\|y\|^2} \right)_+ . \quad \text{minimize } J(\lambda).$$

$$\tilde{\theta} = \hat{\theta}(\Lambda_{const}) = \left(1 - \frac{d\varepsilon^2}{\|y\|^2} \right)_+ y = \hat{\theta}_{st}$$

POSITIVE PART OF STEIN'S ESTI

$\Downarrow$.

$$\boxed{\mathbb{E}_\theta \|\tilde{\theta} - \theta\|^2 \leq \frac{d\varepsilon^2 \|\theta\|^2}{d\varepsilon^2 + \|\theta\|^2} + 4\varepsilon^2.}$$

First oracle inequality

$J(\lambda)$ is SURE (stein unbiased risk estimator)

PF. LEMMA.

$\theta \in \mathbb{R}^d$. $d \geq 4$.

$$\mathbb{E}_\theta \|\hat{\theta}_s - \theta\|^2 \leq \frac{d\varepsilon^2 \|\theta\|^2}{\|\theta\|^2 + d\varepsilon^2} + 4\varepsilon^2.$$

and, $\forall\ d \geq 1$.

$$\mathbb{E}_\theta \|\hat{\theta}_{st} - \theta\|^2 \leq \frac{d\varepsilon^2 \|\theta\|^2}{\|\theta\|^2 + d\varepsilon^2} + 4\varepsilon^2.$$

PF. $\mathbb{E}\|\hat{\theta}_s - \theta\|^2 = d\varepsilon^2 + (-2d(\varepsilon^2 + 4\varepsilon^2 c + c^2) \mathbb{E}_\theta \frac{1}{\|y\|^2}$

$$\mathbb{E}_\theta \frac{1}{\|y\|^2} \geq \frac{1}{\mathbb{E}_\theta \|y\|^2} = \frac{1}{\|\theta\|^2 + \varepsilon^2 d}$$

$$\mathbb{E}_\theta \|\hat{\theta}_s - \theta\|^2 \leq d\varepsilon^2 - (d^2 - 4d)\varepsilon^4 \frac{1}{\|\theta\|^2 + \varepsilon^2 d}$$

$\Rightarrow$ Plug in simplify

Exercise 1.11

Assumption (A)

$$Y_i = f(X_i) + \xi_i.$$

$f: [0,1] \to \mathbb{R}$. ξ_i ind. $\mathbb{E}(\xi_i) = 0$.
$\mathbb{E}(\xi_i^2) = \sigma_\xi^2 < \infty$.

$X_i = \frac{i}{n}$ for $i = 1, \dots, n$.

$\{\varphi_j\}$ $\quad \theta_j = \int_0^1 f \varphi_j \Rightarrow \sum_{j=1}^{\infty} |\theta_j| < \infty$.

SMOOTH SPLINE ESTI

$$f_n^{sp} = \arg\min_{f \in W} \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - f(X_i))^2 + K \int_0^1 (f''(x))^2 dx \right]$$

K: smoothing para.

1) $W: f: [0,1] \to \mathbb{R}$. f' abs cts.

prove f_n^{sp} reproduce polynomials of degree ≤ 1 if $n \ge 2$.

if $f = p$

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - f(X_i))^2 + K \int_0^1 (f''(x))^2 dx$$

$$= \frac{1}{n} \sum_{i=1}^{n} (\xi_i + p(X_i) - f(X_i))^2 + K \int_0^1 (f''(x))^2 dx.$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(\xi_i^2 + (p(X_i) - f(X_i))^2 + \xi_i (p(X_i) - f(X_i))\right) + K \int_0^1 (f''(x))^2 dx$$

ASYMPTOTICALLY EFFICIENT

$$\cancel{\frac{1}{n} \sum_{i=1}^{n} (p(X_i) - f(X_i))^2 + K \int_0^1 (f''(x))^2 dx}.$$

Let $p(x) = ax + b$. $\quad f(x) = cx + d$.

$$\min_{c,d} \frac{1}{n} \sum_{i=1}^{n} (\xi_i + (a-c)X_i + (b-d))^2 = \mathcal{L}(c,d).$$

$$\frac{\partial \mathcal{L}}{\partial c} = -\frac{2}{n} \sum_{i=1}^{n} (\xi_i + (a-c)X_i + (b-d)) X_i + K.$$

Let. $\frac{\partial \mathcal{L}}{\partial c} = 0$ when $(a-c) = 0$.

$$K = \frac{2}{n} \sum_{i=1}^{n} (\xi_i + b - d) X_i$$

$$\frac{\partial \mathcal{L}}{\partial d} = -\frac{1}{n} \sum_{i=1}^{n} (Y_i - cX_i - d) \quad \frac{\partial \mathcal{L}}{\partial d} = 0. \Rightarrow d = \frac{1}{n} \sum_{i=1}^{n} (Y_i - cX_i)$$

$$\cancel{K = \frac{2}{n} \sum_{i=1}^{n} \xi_i + bX_i - dX_i}.$$

$$\frac{\partial \mathcal{L}}{\partial d} = -\frac{1}{n} \sum_{i=1}^{n} (Y_i - cX_i - d) \quad \frac{\partial \mathcal{L}}{\partial d} = 0. \quad d = \frac{1}{n} \sum_{i=1}^{n} (Y_i - cX_i)$$

$$\min_c \frac{1}{n} \sum_{i=1}^{n} \left(\xi_i + (a-c)X_i + b - \left(\frac{1}{n} \sum_{i=1}^{n} \xi_i + aX_i + b - cX_i\right)\right)^2 + Kc = \frac{1}{n} \sum_{i=1}^{n} (\xi_i + aX_i + b - cX_i)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(\xi_i + (a-c)X_i - \left(\frac{1}{n} \sum_{j=1}^{n} \xi_j + a\frac{1}{n} \sum_{j=1}^{n} X_j - c\frac{1}{n} \sum_{j=1}^{n} X_j\right)\right)^2 = \mathcal{L}_2(c)$$

$$\frac{\partial \mathcal{L}_2}{\partial c} = \frac{2}{n} \sum_{i=1}^{n} (\xi_i + \cdots)\left(-X_i + \frac{1}{n} \sum_{j=1}^{n} X_j\right)$$

$$\frac{\partial \mathcal{L}_2}{\partial c} = 0. \quad \cancel{\Rightarrow 2\left(\frac{1}{n} \sum_{i=1}^{n} \left(\xi_i + (a-c)X_i - \frac{1}{n} \sum_{j=1}^{n} \xi_j - a\frac{1}{n} \sum_{j=1}^{n} X_j + \frac{c}{n} \sum_{j=1}^{n} X_j\right)\right)}$$

$$\cancel{Y_i - cX_i - b}.$$

$$2\left(\frac{1}{n} \sum_{i=1}^{n} \left(\xi_i + (a-c)X_i - \frac{1}{n} \sum_{j=1}^{n} \xi_j + \left(\frac{a-c}{n}\right) \sum_{j=1}^{n} X_j\right)\left(\frac{1}{n} \sum_{j=1}^{n} X_j - X_i\right)\right) = 0.$$

if $c = a$. it's zero.

(2) $W = \{f: [0,1] \to \mathbb{R}$ s.t. i) f' abs cts.
ii) $f(0) = f(1)$. $f'(0) = f'(1)$.

$$\min_{\{b_j\}} \sum_{j=1}^{\infty} \left(-2\hat{\theta}_j b_j + b_j^2 (\kappa \pi^4 a_j^2 + 1)[1 + O(n^{-1})] \right)$$

b_j is Fourier coef of f. $O(n^{-1})$ uniformly $\sim \{b_j\}$.

a_j are $W^{per}(\beta, L)$. $\beta = 2$.

Pf consider $W^{per}(\beta, L)$. $W^{per}(\beta, L) \subseteq W$.

$$\frac{1}{n}\sum_{i=1}^{n}\Big(Y_i - \sum_{j=1}^{\infty} b_j\varphi_j(X_i)\Big)^2 + \kappa\int_0^1 (f''(x))^2 dx$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big(Y_i^2 - 2\sum_{j=1}^{\infty} Y_i b_j \varphi_j(X_i) + \Big[\sum_{j=1}^{\infty} b_j\varphi_j(X_i)\Big]^2\Big) + \kappa\int_0^1 (f''(x))^2 dx$$

$$\Leftrightarrow -\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{\infty} Y_i b_j \varphi_j(X_i) + \Big[\sum_{j=1}^{\infty} b_j \varphi_j(X_i)\Big]^2 + \kappa\int_0^1 f''^2$$

NOTE This is particularly for tricangular basis.

even j. $\frac{d^2}{dx^2}\varphi_j(x) = \frac{d^2}{dx^2}\cos\left(2\pi\frac{j}{2}x\right) = -\pi^2 j^2 \varphi_j(x)$.

odd j. $\frac{d^2}{dx^2}\varphi_j(x) = \frac{d^2}{dx^2}\sin\left(2\pi\frac{j-1}{2}x\right) = -\pi^2 (j-1)^2 \varphi_j(x)$.

$$\frac{d^2}{dx^2}\varphi_j(x) = -\pi^2 a_j \varphi_j(x)$$

$$\int_0^1 (f''(x))^2 dx = \sum_{j=1}^{\infty}\left(-\pi^2 a_j b_j\right)^2$$

$$-2\sum_{i=1}^{n} b_i\hat{\theta}_i + \sum_{j=1}^{\infty}\pi^4 a_j^2 b_j^2 + \frac{1}{n}\sum_{i=1}^{n} f(X_i)^2$$

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\sum_{j=1}^{\infty} b_j\varphi_j(X_i)\Big)^2 - \sum_{j=1}^{\infty} b_j^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} b_j b_k \varphi_j(X_i)\varphi_k(X_i) - \sum_{j=1}^{\infty} b_j^2.$$

$$= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} b_j b_k \frac{1}{n}\sum_{i=1}^{n}\varphi_j(X_i)\varphi_k(X_i) - \sum_{j=1}^{\infty} b_j^2.$$

$\frac{a^2+b^2}{2} \geq ab$.

$$= \cancel{\sum_{j=1}^{\infty} b_j^2\Big(\frac{1}{n}\sum_{i=1}^{n}\varphi_j(X_i)^2 - 1\Big) - \sum_{j=1}^{\infty} b_j}$$

$$\cancel{\frac{1}{n}\sum_{i=1}^{n}\varphi_j(X_i)\varphi_k(X_i)} \to \delta_{jk}(1 + O(n^{-1}))$$

$$\cancel{\leq \sum_{j=1}^{\infty}\Big(\sum_{k=1}^{\infty} b_k\Big)^2 \frac{1}{n}\Big(\sum_{i=1}^{n}\varphi_j(X_i)^2\Big)}$$

$$\frac{1}{n}\sum_{i=1}^{n}\varphi_j(X_i)\varphi_k(X_i) = \cancel{\int_0^1 \varphi_j(x)\varphi_k(x)dx} = \delta_{jk}. \quad \text{for } 1 \leq j, k \leq n$$

$$= \sum_{j=1}^{n-1}\sum_{k=1}^{n-1} b_j b_k \delta_{jk} - \sum_{j=1}^{\infty} b_j^2$$

$$\sum_{j=n}^{\infty}\sum_{k=n}^{\infty} b_j b_k$$

(3) $\min_{\{b_j\}} \sum_{j=1}^{\infty} \left(-2\hat{\theta}_j b_j + b_j^2(K\pi^4 a_j^2+1)\right) \cancel{[1+O(n^{-1})]}$

$\underbrace{\qquad}_{\mathcal{L}(b_j)}$

$\frac{\partial f}{\partial b_j} = \sum_{j=1}^{\infty} \left(-2\hat{\theta}_j + 2b_j(K\pi^4 a_j^2+1)\right).$

$b_j^* = \frac{\hat{\theta}_j}{K\pi^4 a_j^2+1}$ $\lambda_j = \frac{1}{K\pi^4 a_j^2+1}$

Example. 3.2. Projection Estimator. (C_p).

$\Lambda_{proj} = \{\lambda \mid \lambda_j = I\{j<N\},\ N\in\{1,2,\cdots,N_{max}\}\}.$

$$\hat{\theta}_{j,N} = \begin{cases} y_j & \text{if } 1\le j\le N. \\ 0 & \text{if } j>N. \end{cases}$$

$$J(\lambda) = \sum_{j\le N}\left(y_j^2 - 2(y_j^2-\varepsilon^2)\right) = 2N\varepsilon^2 - \sum_{j\le N} y_j^2.$$

$\hat{\lambda}_j = I\{j\le \tilde{N}\}.$

$\hat{N} = \operatorname*{argmin}_{1\le N\le N_{max}} \left(2N\varepsilon^2 - \sum_{j\le N} y_j^2\right).$

$\tilde{N} = \operatorname*{arg\,min}_{1\le N\le N_{max}} \left\{\sum_{j=1}^{N_{max}} (y_j - \hat{\theta}_{j,N})^2 + 2N\varepsilon^2\right\}.$

C_p -criterion with $\varepsilon = \frac{\sigma}{\sqrt{n}}.$

EX 3.3.

$\Lambda_{spline} = \{\lambda \mid \lambda_j = \frac{1}{1+s a_j^2} I\{j\le N_{max}\}, s\in S,\ \beta\in B\}$

$S \subseteq (0,\infty),\ B\subseteq (0,\infty).$

$N_{max} = \lfloor \frac{1}{\varepsilon^2} \rfloor$

corresponding non-linear esti $\tilde{\theta}$.

has $\hat{\lambda}(\Lambda_{spline})$. mini $J(\lambda)$.

EX 3.4.

$$\Lambda_{pinsker} = \{\lambda \mid \lambda_j = (1-s a_j)_+ I\{j \le N_{max}\}, s \in S, \beta \in B\}$$

Ex 3.5. monotone ~~class~~ weights class

$$\Lambda_{mon} = \{\lambda \mid 1 \ge \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{N_{max}} \ge 0,\ \lambda_j = 0, j > N_{max}\}.$$

monotone oracle.

$$\lambda^{oracle}(\Lambda_{mon}, \theta) = \arg\min_{\lambda \in \Lambda_{mon}} R(\lambda, \theta)$$

$$\tilde{\lambda} = \tilde{\lambda}(\Lambda_{mon}) = \arg\min_{\lambda \in \Lambda_{mon}} J(\lambda)$$

Ex. 3.6. blockwise constant weights

$$\{1, 2, \cdots, N_{max}\} \to \bigcup_{j=1}^{J} B_j = \{1, 2, \cdots, N_{max}\}.$$

$$B_i \cap B_j = \emptyset,\ i \ne j.$$

if $\min\{k : k \in B_j\} > \max\{k : k \in B_{j-1}\}$

$$\Lambda^* = \{\lambda \mid \lambda_k = \sum_{j=1}^{J} t_j I(k \in B_j) : 0 \le t_j \le 1, j = 1, \cdots, J\}.$$

$$\tilde{\lambda} = \arg\min_{\lambda \in \Lambda^*} J(\lambda) = \sum_{j=1}^{J} \tilde{\lambda}_{(j)} I(k \in B_j)$$

$$\tilde{\lambda}_{(j)} = \arg\min_{t \in [0,1]} \sum_{k \in B_j} \left(t^2 y_k^2 - 2t(y_k^2 - \varepsilon^2)\right).$$

$$= \left(1 - \frac{\varepsilon^2 \operatorname{Card} B_j}{\sum_{k \in B_j} y_k^2}\right)_+$$

$$\tilde{\theta}_k = \begin{cases} \tilde{\lambda}_{(j)} y_k & k \in B_j, j = 1, 2, \cdots, J \\ 0 & k > N_{max} \end{cases}$$

Blockwise constant positive part. Stein estimator.

ORACLE inequality of Block Stein esti.

by first oracle inequality. for each block B_j.

$$\mathbb{E}_\theta \|\tilde{\theta}_{(j)} - \theta\|_{(j)}^2 \le \min_{t_j} \sum_{k \in B_j} [(1-t_j)^2 \theta_k^2 + \varepsilon^2 t_j^2] + 4\varepsilon \quad j = 1, 2,$$

$$\Rightarrow \mathbb{E}_\theta \|\tilde{\theta} - \theta\|^2 = \sum_{j=1}^{J} \mathbb{E}_\theta \|\tilde{\theta}_{(j)} - \theta_{(j)}\|_{(j)}^2 + \sum_{k > N_{max}} \theta_k^2$$

$$\le \sum_{j=1}^{J} \min_{t_j} \sum_{k \in B_j} [(1-t_j)^2 \theta_k^2 + \varepsilon^2 t_j^2] + \sum_{k > N_{max}} \theta_k^2 + 4J\varepsilon^2$$

$$= \underbrace{\min_{\lambda \in \Lambda^*} R(\lambda, \theta) + 4J\varepsilon^2}_{\text{second oracle inequality}}.$$

① non-asymptotic

② mimics the blockwise constant oracle.

$$\lambda^{oracle}(\Lambda^*,\theta)=\arg\min_{\lambda\in\Lambda^*} R(\lambda,\theta).$$

Assumption (C). $\exists \eta>0$. s.t.

$$\max_{1\le j\le J-1}\frac{T_{j+1}}{T_j}\le 1+\eta$$

$\Downarrow$

Lemma.

if (C). $\forall\theta\in\ell^2(\mathbb{N})$.

$$\Rightarrow \min_{\lambda\in\Lambda^*} R(\lambda,\theta)\le(1+\eta)\min_{\lambda\in\Lambda_{mon}} R(\lambda,\theta)+\varepsilon^2 T_1$$

Pf: Goal $\begin{cases}\forall\lambda\in\Lambda_{mon}.\ \exists\bar\lambda\in\Lambda^* \text{ s.t.}\\ R(\bar\lambda,\theta)\le(1+\eta)R(\lambda,\theta)+\varepsilon^2T_1,\ \forall\theta\in\ell^2(\mathbb{N})\end{cases}$

Let. $\bar\lambda_k=\begin{cases}\bar\lambda_{(j)}\triangleq\max_{m\in B_j}\lambda_m & \text{if } k\in B_j.\\ 0, & \text{if } k>N_{max}\end{cases}$

1. $\bar\lambda_k\ge\lambda_k$, $\forall k$.

$$\Rightarrow R(\bar\lambda,\theta)=\sum_{k=1}^\infty\left[(1-\bar\lambda_k)^2\theta_k^2+\varepsilon^2\bar\lambda_k^2\right]\le\sum_{k=1}^\infty\left[(1-\lambda_k)^2\theta_k^2+\varepsilon^2\bar\lambda_k^2\right]$$

$$R(\lambda,\theta)=\sum_{k=1}^\infty\left[(1-\lambda_k)^2\theta_k^2+\varepsilon^2\lambda_k^2\right].$$

it suffices to prove

$$\varepsilon^2\sum_{k=1}^\infty\bar\lambda_k^2\le(1+\eta)\,\varepsilon^2\sum_{k=1}^\infty\lambda_k^2+\varepsilon^2T_1$$

$$\sum_{k=1}^\infty\bar\lambda_k^2=\sum_{k\le N_{max}}\bar\lambda_k^2$$

$$\le T_1+\sum_{j=2}^J\sum_{k\in B_j}\bar\lambda_k^2\quad(0\le\bar\lambda_1\le1).$$

$$=T_1+\sum_{j=2}^J T_j\,\bar\lambda_{(j)}^2$$

$$\le T_1+(1+\eta)\sum_{j=2}^J T_{j-1}\bar\lambda_{(j)}^2$$

$$\le T_1+(1+\eta)\sum_{j=2}^J\sum_{m\in B_{j-1}}\lambda_m^2\quad(\bar\lambda_{(j)}\le\lambda_m,\ \forall m\in B_{j-1},\ \text{monotone class.})$$

$$=T_1+(1+\eta)\sum_{j=1}^{J-1}\sum_{m\in B_j}\lambda_m^2$$

$$\le T_1+(1+\eta)\sum_{k=1}^\infty\lambda_k^2$$

$\square$.

$\Rightarrow$ Th!. $\tilde\theta$: block stein esti $\tilde\theta$

$$\mathbb{E}_\theta\|\tilde\theta-\theta\|^2\le(1+\eta)\min_{\lambda\in\Lambda_{mon}}R(\lambda,\theta)+\varepsilon^2(T_1+4J)$$

① How to construct good systems of blocks. i.e.. $\{B_j\}$ s.t. η and $\varepsilon^2(T_1+4J)$ are sma

Choose block

$$T_1 = \lceil \rho_\varepsilon \rceil$$

$$T_2 = \lfloor T_1 (1+\rho_\varepsilon) \rfloor$$

$$\vdots$$

$$T_{J-1} = \lfloor T_1 (1+\rho_\varepsilon)^{J-2} \rfloor$$

$$T_J = N_{max} - \sum_{j=1}^{J-1} T_j$$

$$J = \min\{m : T_1 + \sum_{j=2}^{m} \lfloor T_1 (1+\rho_\varepsilon)^{j-1} \rfloor \geq N_{max}\}.$$

$$\rho_\varepsilon = \left(\log\left(\tfrac{1}{\varepsilon}\right)\right)^{-1}$$

THIRD ORACLE INEQ.

COR

$$\boxed{E_\theta \|\tilde\theta - \theta\|^2 \leq (1+3\rho_\varepsilon) \min_{\lambda \in \Lambda_{mon}} R(\lambda, \theta) + C\varepsilon^2 \log^2(1/\varepsilon).}$$

$\forall\, \theta \in \ell^2(\mathbb{N})$. $\forall\, 0 < \varepsilon < \varepsilon_0$

since $\rho_\varepsilon = o(1)$. $\Rightarrow$ oracle is exactly adapted.

It still holds to class. $\Lambda = \Lambda_{proj}, \Lambda_{spline}, \Lambda_{pinsk}$.

for any class $\Lambda \subset \Lambda_{mono}$

Def. ASYMPTOTICALLY EFFICIENT.

$$\lim_{\varepsilon \to 0} \frac{\sup_{\theta \in \Theta} E_\theta \|\theta^*_\varepsilon - \theta\|^2}{\inf_{\hat\theta_\varepsilon} \sup_{\theta \in \Theta} E_\theta \|\hat\theta_\varepsilon - \theta\|^2} = 1.$$

$\inf$ over all estimators

$\theta \in \Theta(\beta, Q)$. Gaussian sequential model

① Simplified Pinsker estimator $\hat\theta(\ell^*)$.

and Pinsker estimator $\hat\theta(\ell)$.

are asymptotically on the class $\Theta(\beta, Q)$.

② But they depends on β, Q unknown.

Def. ADAPTIVE IN THE EXACT MINIMAX SENSE

θ^*_ε is asymptotically efficient for all classes $\Theta(\beta, Q)$, $\beta > 0$, $Q > 0$. simultaneously

TH. Stein WGB esti $\tilde\theta$

is adaptive in the exact minimax sense on family of Sobolev ellipsoid

$$\{\Theta(\beta, Q), \beta > 0, Q > 0\}$$

GAUSSIAN SEQUENTIAL

$y_j = \theta_j + \varepsilon \xi_j \quad . \; j = 1, 2, \cdots,$

$\theta = (\theta_1, \cdots) \in \ell^2(\mathbb{N})$. $\quad 0 < \varepsilon < 1$

iid. $\mathcal{N}(0,1)$.

$y \leftarrow$ OBSERVATION.

$\boxed{\text{PINSKER ESTIMATOR} \quad \text{Tsy. BOOK}}$

RISK.

$R(\lambda, \theta) = \mathbb{E}_f \| f_{\varepsilon,\lambda} - f \|_2^2 = \mathbb{E}_\theta \left[\| \hat{\theta}(\lambda) - \theta \|^2 \right]$

$\hat{\theta}_j(\lambda) = \lambda_j y_j \quad j = 1, 2, \cdots$

$\hat{\Theta}(\lambda) = (\hat{\theta}_1(\lambda), \cdots,$

$R(\lambda, \theta) = \sum_{j=1}^{\infty} \left[(1-\lambda_j)^2 \theta_j^2 + \varepsilon^2 \lambda_j^2 \right]$

LINEAR ESTIMATORS.

$\lambda_j \in \mathbb{R}$ WEIGHTS.

$\lambda \in \ell^2(\mathbb{N})$.

$f_{\varepsilon;\lambda}(x) = \sum_{j=1}^{\infty} \lambda_j y_j \varphi_j(x)$. $\quad ☆$

E.g. $\hat{f}_\varepsilon(x) = \sum_{j=1}^{N} \ell_j^* y_j \varphi_j(x)$ PINSKER.

$\ell_j^* = (1 - \kappa^* a_j)_+ \qquad N_\varepsilon = \max\{j : \ell_j^* > 0\}$.

$\kappa^* = \left(\frac{\beta}{(2\beta+1)(\beta+1)Q} \right)^{\beta/(2\beta+1)} \varepsilon^{\frac{2\beta}{2\beta+1}} \qquad N_\varepsilon \to \infty \; (\varepsilon \to 0)$.

$a_j = \begin{cases} j^\beta, & j \text{ even} \\ (j-1)^\beta, & j \text{ odd}. \end{cases}$

$☆ : R(\lambda, \theta) = \sum_{j=1}^{\infty} \mathbb{E}_f \left[(\lambda_j y_j - \theta_j)^2 \right]$.

$= \sum_{j=1}^{\infty} \mathbb{E}(\lambda_j(\theta_j + \varepsilon \xi_j) - \theta_j)^2 = \sum_{j=1}^{\infty} \mathbb{E}\left[(\lambda_j - 1)\theta_j + \varepsilon \xi_j \lambda_j \right]^2$

LINEAR MINIMAX.

$\inf_\lambda \sup_{f \in \tilde{W}(\beta, L)} \mathbb{E}_f \| f_{\varepsilon,\lambda} - f \|_2^2 = \inf_\lambda \sup_{\theta \in \Theta(\beta, Q)} \mathbb{E}_\theta \| \hat{\theta}(\lambda) - \theta \|^2$

$= C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1 + o(1)) \quad \varepsilon \to 0$.

PINSKER CONSTANT.

$\boxed{\#: \frac{\varepsilon^2}{\kappa} \sum_{j=1}^{\infty} a_j (1 - \kappa a_j)_+ = Q \quad \text{ellipsoid constant}}$

solution exist $\kappa = \kappa(\varepsilon) > 0$

unique: $\kappa = \frac{\varepsilon^2 \sum_{m=1}^{N} a_m}{Q + \sum_{m=1}^{N} \varepsilon^2 a_m^2} \qquad a_j \uparrow \to +\infty$.

$N = \max\{ j : \varepsilon^2 \sum_{m=1}^{j} a_m (a_j - a_m) < Q \} < +\infty$.

PF. $\mathcal{N}$ is finite ① $\forall j \le N$: $\qquad$ Analyze diff.

$Q > \varepsilon^2 \sum_{m=1}^{N} a_m (a_N - a_m) \geq \varepsilon^2 \sum_{m=1}^{N} a_m (a_j - a_m)$.

② $\forall j > N$: $\geq \varepsilon^2 \sum_{m=1}^{N} a_m (a_{N+1} - a_m) = \varepsilon^2 \sum_{m=1}^{N+1} a_m (a_{N+1} - a_m) \geq Q$.

$\Rightarrow \quad 1 > \kappa a_j$ for $j \le N$; $1 \le \kappa a_j$ for $j > N$.

$N = \max\{j : a_j < \frac{1}{\kappa}\}$.

$\Rightarrow \frac{\varepsilon^2}{\kappa} \sum_{j=1}^{\infty} a_j (1-\kappa a_j)_+ = \frac{\varepsilon^2}{\kappa} \sum_{j=1}^{N} a_j (1-\kappa a_j) = Q$. $\square$

$l_j := (1-\kappa a_j)_+ \quad j=1,2,\cdots \quad l = (l_1, l_2, \cdots)$;

$D^* := \varepsilon^2 \sum_{j=1}^{\infty} (1-\kappa a_j)_+ = \varepsilon^2 \sum_{j=1}^{\infty} l_j$

Eq has solution. $\Leftarrow$ positive part. $\not\Leftarrow$ analyze positive

$\varepsilon^2 \sum_{j=1}^{\infty} a_j (1-\kappa a_j)_+ = Q\kappa$. $\quad a_j \to \infty$.

$\|$

$\varepsilon^2 \sum_{j=1}^{N} a_j (1-\kappa a_j)_+ = Q\kappa$. $\quad N := \max\{j : 1 \geq \kappa a_j\}$,

$\frac{Q}{\varepsilon^2} \kappa = \sum_{j=1}^{N} a_j - \kappa \sum_{j=1}^{N} a_j^2$

$$\kappa = \frac{\varepsilon^2 \sum_{j=1}^{N} a_j}{Q + \varepsilon^2 \sum_{j=1}^{N} a_j^2}$$

~~$N := \max\{j : 1 > \varepsilon^2 \sum_{j=1}^{N} a_j\}$~~

UNIQE

$N := \{j : 1 > \varepsilon^2 \sum_{j=1}^{j}$

$$\kappa = \frac{\varepsilon^2 \sum_{j=1}^{\max\{j : 1 > \kappa a_j\}} a_j}{Q + \varepsilon^2 \sum_{j=1}^{\max\cdots} a_j^2}.$$

(n) : Ellipsoid. $\sum_{j=1}^{\infty} \theta_j^2 a_j^2 \leq Q$. $\quad a_j \geq 0$. s.t. $\mathrm{Card}\{j : a_j = 0\} < \infty$.

κ exists for #. $\quad D^* < \infty$.

$\Rightarrow$. $\inf_{\lambda} \sup_{(n)} R(\lambda, \theta) = \sup_{\theta \in (n)} \inf_{\lambda} R(\lambda, \theta) = \sup_{\theta \in \Theta} R(l, \theta) = D^*$.

MINIMAX OF REAL PINSKER.

PF. $\sup_{\theta \in (n)} \inf_{\lambda} R(\lambda, \theta) \leq \inf_{\lambda} \sup_{\theta \in (n)} R(\lambda, \theta) \leq \sup_{\theta \in \Theta} R(l, \theta)$.

① . $\sup_{\theta \in (n)} R(l, \theta) \leq D^*$.

② $\sup_{\theta \in (n)} \inf_{\lambda} R(\lambda, \theta) \geq D^*$.

①: $R(l, \theta) = \sum_{j=1}^{\infty} \left((1-l_i)^2 \theta_i^2 + \varepsilon^2 l_i^2 \right)$

$= \varepsilon^2 \sum_{j=1}^{\infty} l_j^2 + \underline{\sum_{j=1}^{\infty} (1-l_j)^2 \theta_j^2} \; \sum_{j=1}^{\infty} (1-(1-\kappa a_j)_+)^2 \theta_j^2$.

$= \varepsilon^2 \sum_{j=1}^{\infty} l_j^2 + \sum_{j=1}^{\infty} (1-l_j)^2 a_j^{-2} a_j^2 \theta_j^2$

$\leq \varepsilon^2 \sum_{j=1}^{\infty} l_j^2 + Q \underline{\sup_{i : a_i > 0} [(1-l_i)^2 a_j^{-2}]} \leq \varepsilon^2 \sum_{j=1}^{\infty} l_j^2 + Q\kappa^2$

$1 - \kappa a_j \leq l_j \leq 1$. $\quad \sup_{a_j > 0} \left(\frac{1-l_j}{a_j}\right)^2 = \sup_{a_j > 0} \left(\frac{1-(1-\kappa a_j)_+}{a_j}\right)^2$.

$= \kappa^2$.

$= \varepsilon^2 \sum_{j=1}^{\infty} l_j^2 + \underline{\varepsilon^2 \kappa \sum_{j=1}^{\infty} a_j l_j}$ by #.

$= \varepsilon^2 \sum_{j=1}^{\infty} l_j (l_j + \kappa a_j) = \varepsilon^2 \sum_{j=1}^{\infty} (1-\kappa a_j)_+ ((1-\kappa a_j)_+ + \kappa a_j)$

$$\boxed{= \varepsilon^2 \sum_{j : l_j \neq 0} l_j = D^*.} \; \square$$

$$V := \{v : v_j \in \mathbb{R},\ v_j^2 = \frac{\varepsilon^2(1-\kappa a_j)_+}{\kappa a_j} \text{ if } a_j > 0, \quad v_j \text{ if } a_j = 0\}.$$

$\Rightarrow$ By #. $V \subset \Theta$:$\Leftrightarrow$. $\sum_{j=1}^{\infty} a_j^2 v_j^2 = Q$.

$\Rightarrow$.

$$\sup_{\Theta \in \Theta} \inf_{\lambda} R(\lambda, \theta) \geq \sup_{v \in V} \inf_{\lambda} \sum_{i=1}^{\infty} [(1-\lambda_i)^2 v_i^2 + \varepsilon^2 \lambda_i^2]$$

Optimize in λ.

$$= \sup_{v \in V} \left[\sum_{i: a_i = 0} \frac{v_i^2 \varepsilon^2}{v_i^2 + \varepsilon^2} + \sum_{i: a_i > 0} \frac{v_i^2 \varepsilon^2}{v_i^2 + \varepsilon^2} \right]$$

$$= \varepsilon^2 \operatorname{Card}\{i : a_i = 0\} + \sum_{i: a_i > 0} \frac{\varepsilon^4 (1-\kappa a_i)_+}{\kappa a_i \left(\frac{\varepsilon^2(1-\kappa a_i)_+}{\kappa a_i} + \varepsilon^2\right)}.$$

$$= \varepsilon^2 \operatorname{Card}\{i : a_i = 0\} + \sum_{i: a_i > 0} \frac{\varepsilon^4 (1-\kappa a_i)_+}{\varepsilon^2(1-\kappa a_i)_+ + \varepsilon^2 \kappa a_i}$$

$$= \varepsilon^2 \operatorname{Card}\{i : a_i = 0\} + \varepsilon^2 \sum_{i: a_i > 0} (1-\kappa a_i)_+$$

$$= \varepsilon^2 \sum_{i=1}^{\infty} (1-\kappa a_i)_+ = D^* \qquad \square.$$

PINSKER ESTIMATOR for general Θ.

$$\Theta = \{\theta = \{\theta_j\} : \sum_{j=1}^{\infty} a_j^2 \theta_j^2 \leq Q\}.$$

$\kappa(\varepsilon) :=$ solution of $\frac{\varepsilon^2}{\kappa} \sum_{j=1}^{\infty} a_j (1-\kappa a_j)_+ = Q$

if $a_j \to \infty \uparrow \Rightarrow \kappa = \frac{\varepsilon^2 \sum_{m=1}^{N} a_m}{Q + \sum_{m=1}^{N} \varepsilon^2 a_m^2}$ ~~$N:=\{j:\}$~~

$\xrightarrow{\text{solution exists}}$ $\ell_j \triangleq (1-\kappa a_j)_+ \quad j = 1, 2, \ldots$

$$\hat{f}_\varepsilon(x) = \sum_{j=1}^{\infty} \ell_j^* y_j \varphi_j(x).$$

PINSKER estimator is trying to choose weights of linear estimator to minimize the risk? while keeping the estimator itself inside ellipsoide?

How ~~does~~ did anyone came up with # stupid equation ???

Doesn't make sense at all !!!

~~$\rho(x,y)=\rho(Tx,Ty)$~~ 1. Is PK-e inside Ⓜ?

with #: $\sum l_i^2 a_i^2 = \sum a_i^2 (1-Ka_i)_+^2 \approx Q$.

2. Where the fucking hell did # come from?

LEMMA

Ⓜ = Ⓜ(β, Q). $\varepsilon > 0$.

$a_j = \begin{cases} j^\beta & \text{even } j \\ (j-1)^\beta & \text{odd } j \end{cases}$.

$\Rightarrow$ (i). $\exists K$ solution of #. unique and

$K = K^*(1+o(1))$ as $\varepsilon \to 0$.

$K^* = \left(\frac{\beta}{(2\beta+1)(\beta+1)Q}\right)^{\frac{\beta}{2\beta+1}} \varepsilon^{\frac{2\beta}{2\beta+1}}$

(ii). $D^* = C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1+o(1))$. $\varepsilon \to 0$.

$C^* = [Q(2\beta+1)]^{\frac{1}{2\beta+1}} \left(\frac{\beta}{\beta+1}\right)^{\frac{2\beta}{2\beta+1}}$

(iii) $\max_{j \geq 2} l_j^2 a_j^2 = O(\varepsilon^{\frac{2}{2\beta+1}})$ $\varepsilon \to 0$.

REAL and SIMPLIFIED PINSKER ARE SIMILAR.

PF: Existence of solution is known. uniqueness done.

$\frac{\varepsilon^2}{K} \sum_{j=1}^\infty a_j (1-Ka_j)_+ = Q$

$= \frac{2\varepsilon^2}{K} \sum_{m=1}^{M} (2m)^\beta (1-K(2m)^\beta)$

$M = \lfloor (1/K)^{1/\beta} / 2 \rfloor$

$\Rightarrow \sum_{m=1}^{M} m^a = \frac{M^{a+1}}{a+1}(1+o(1))$ $M \to \infty$.

~~$\Rightarrow Q = \frac{2\varepsilon^2}{K} \sum_{m=1}^{M} (2m)^\beta - K(2m)^{2\beta}$~~

$\frac{2\varepsilon^2}{K} \sum_{m=1}^{M} (2m)^\beta (1-K(2m)^\beta) \overset{?}{=} \frac{2\varepsilon^2}{K} \sum_{m=1}^{\infty} (2m)(1-K(2m))^\beta_+$

Only finite elements remains.

$1 - K(2m)^\beta < 0 \Leftrightarrow \frac{1}{2}\left(\frac{1}{K}\right)^{1/\beta} = K$.

$Q = \frac{2\varepsilon^2}{K} \sum_{m=1}^{\infty} (2m)^\beta - 2\varepsilon^2 \sum_{m=1}^{M} (2m)^{2\beta}$

$= \frac{2\varepsilon^2}{K} 2^\beta \sum_{m=1}^{\infty} m^\beta - 2^{2\beta+1} \varepsilon^2 \sum_{m=1}^{M} m^{2\beta}$

$= \frac{2^{\beta+1}\varepsilon^2}{K} \frac{M^{\beta+1}}{\beta+1}(1+o(1)) - 2^{2\beta+1} \varepsilon^2 \frac{M^{2\beta+1}}{2\beta+1}(1+o(1))$

$= \frac{2^{\beta+1} \varepsilon^2 M^{\beta+1}}{K(\beta+1)(2\beta+1)} \left(2\beta+1 + \overset{o(1)}{\cancel{o(2\beta+1)}} - 2^\beta(\beta+1)M^\beta K (1+o(1))\right)$

~~$= 2^{\beta+1} \varepsilon^2 M^{\beta+1} (2\beta+1 - 2^{\beta+1}$~~ $- 2^\beta$ plugin M. Done

(ii) $\underline{D^* = C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1 + o(1))}$ $\varepsilon \to 0$

$D^* = \varepsilon^2 \sum_{j=1}^{\infty} (1 - Ka_j)_+$ same.

(iii). $\forall j > N$. $\nu_j^2 = 0$.

$N = \max\{j : \varepsilon^2 \sum_{m=1}^{j} a_m (a_j - a_m) < Q\}$

$\nu_j^2 = \frac{\varepsilon^2 (1 - Ka_j)_+}{Ka_j}$ $K = \frac{\varepsilon^2 \sum_{m=1}^{N} a_m}{Q + \sum_{m=1}^{N} \varepsilon^2 a_m^2}$

$1 - Ka_j > 0 \Rightarrow 1 > \frac{\varepsilon^2 \sum_{m=1}^{N} a_m}{Q + \sum_{m=1}^{N} \varepsilon^2 a_m^2} \Rightarrow Q > \varepsilon^2 \sum_{m=1}^{N} (a_m - a_m^2)$ ✓.

$a_N < \frac{1}{K}$.

$\nu_j^2 a_j^2 = \frac{\varepsilon^2 a_j (1 - Ka_j)_+}{K} \leq \frac{\varepsilon^2 a_N}{K} \leq \frac{\varepsilon^2}{K^2} = O(\varepsilon^{\frac{2}{\beta+1}})$.

COR. ☆ ASYMPTOTICS FOR RISK OF LINEAR ESTIMATOR

$\hat{\Theta}(\ell)$. Pinsker. $\Theta(\beta, Q)$ ellipsoid

$\Rightarrow \inf_{\lambda} \sup_{\theta \in \Theta(\beta, Q)} \mathbb{E}_\theta \| \hat{\theta}(\lambda) - \theta \|^2 = \sup_{\theta \in \Theta(\beta, Q)} \mathbb{E}_\theta \| \hat{\theta}(\ell) - \theta \|^2$.

$= C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1 + o(1))$.

value of linear lower bound.

$\varepsilon \to 0$. C^* Pinsker const.

TH. PINSKER.'s (Gaussian sequential)

$\beta > 0$ $L > 0$.

$\Rightarrow \lim_{\varepsilon \to 0} \sup_{f \in \tilde{W}(\beta, L)} \varepsilon^{-\frac{4\beta}{2\beta+1}} \mathbb{E}_f \| \hat{f}_\varepsilon - f \|_2^2 = \lim_{\varepsilon \to 0} \inf_{T_\varepsilon} \sup_{f \in \tilde{W}(\beta, L)} \varepsilon^{-\frac{4\beta}{2\beta+1}} \mathbb{E}_f \| T_\varepsilon - f \|_2^2$

$= C^*$.

$\inf T_\varepsilon$ all esti, $\mathbb{E}_f$: model. $\hat{f}_\varepsilon = \sum_{j=1}^{\infty} \ell_j^* y_j \varphi_j$

Pf: ① Upper bound. of Risk of $\hat{f}_\varepsilon$ to f. $\ell_j^* = (1 - K^* a_j)_+$.

$\sup_{f \in \tilde{W}(\beta, L)} \mathbb{E}_f \| \hat{f}_\varepsilon - f \|_2^2 \leq C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1 + o(1))$, $\varepsilon \to 0$. $K^* = (\ldots)$

② lower bound on minimax

$R_\varepsilon^* \triangleq \inf_{T_\varepsilon} \sup_{f \in \tilde{W}(\beta, L)} \mathbb{E}_f \| T_\varepsilon - f \|_2^2 \geq$ ——

PRELIMINARIES.

BAYES. in 1 dim $x = a + \varepsilon \xi$ $a \in \mathbb{R}$. $\xi \sim N(0,1)$ $\varepsilon > 0$.

$\hat{a} = \hat{a}(x)$. risk: $\mathbb{E}(\hat{a} - a)^2$

BAYES RISK. to $N(0, s^2)$ $R^B(\hat{a}) = \int \mathbb{E}[(\hat{a} - a)^2] \mu_s(a) da$.

$= \int_{\mathbb{R}^2} (\hat{a}(x) - a)^2 \mu_\varepsilon(x - a) \mu_s(a) dx da$

$\mu_s(u) = \frac{1}{s} \varphi(\frac{u}{s})$. → density of N

$\hat{a}^B := \arg\min_{a} R^B(\hat{a})$.

$R^B(\hat{a}) = \mathbb{E}[(\hat{a}(x) - a)^2]$. $\mathbb{E}$: to Gaussian pair (x, a) indep

$\hat{a}^B = \mathbb{E}(a|x)$. $R^B(\hat{a}^B) = \min_{\hat{a}} R^B(\hat{a}) = \mathbb{E}[\mathrm{Var}(a|x)]$.

$a = \frac{S^2}{\varepsilon^2 + S^2} x$. $= \frac{S^2 \varepsilon^2}{\varepsilon^2 + S^2}$

posterior distribution.

① Upper bound.

$$\mathbb{E}_f\|\hat{f}_\varepsilon - f\|_2^2 = \mathbb{E}_\theta\|\hat{\theta}(\ell^*) - \theta\|^2 = R(\ell^*, \theta).$$ L^2 risk $\to$ ℓ^2 risk of coef.

$$R(\ell^*, \theta) = \mathbb{E}\sum_{j=1}^{\infty}(\theta_j - \ell_j^* y_j)^2 = \mathbb{E}\sum_{j=1}^{\infty}(\theta_j - (1-\kappa^* a_j)_+ y_j)^2$$

$$= \mathbb{E}\sum_{j=1}^{\infty}(\theta_j - \ell_j^*(\theta_j + \varepsilon\xi_j))^2 = \mathbb{E}\sum_{j=1}^{\infty}((1-\ell_j^*)\theta_j - \varepsilon\xi_j\ell_j^*)^2.$$

$$= \sum_{j=1}^{\infty}(1-\ell_j^*)^2\theta_j^2 + \sum_{j=1}^{\infty}\varepsilon^2\ell_j^{*2} = \varepsilon^2\sum_{j=1}^{\infty}\ell_j^* + \sum_{j=1}^{\infty}(1-(1-\kappa^* a_j)_+)^2\theta_j^2$$ ← Try to get. Q.

$$\cdots \le \varepsilon^2\sum_{j=1}^{\infty}(\ell_j^*)^2 + Q(\kappa^*)^2.$$

Let $M^* = \lfloor (\frac{1}{\kappa^*})^{1/\beta}/2 \rfloor$, $M = \lfloor (1/\kappa)^{1/\beta}/2 \rfloor$ κ solution of #.

only M^* terms are non-zero.

$$= \varepsilon^2\left(\sum_{j=1}^{M}(\ell_j^*)^2\right) + Q(\kappa^*)^2 = \varepsilon^2\left(\sum_{j=1}^{M}(1-\kappa^* a_j)^2\right) + Q\kappa^2(1+o(1))$$

= Standard computation.

② Lower bound. ~~$X = a + \varepsilon\xi$, $a \in \mathbb{R}$, $\xi \sim N(0,1)$, $\varepsilon > 0$~~.

1. Reduction to parametric family.

$N = \max\{j : \ell_j > 0\}$ ℓ_j Pinsker weight

$$\Theta_N = \{\theta^N = (\theta_2, \cdots, \theta_N) \in \mathbb{R}^{N-1} : \sum_{j=2}^{N} a_j^2\theta_j^2 < Q\}.$$

Parametric $\mathcal{F}_N = \{f_{\theta^N}(x) = \sum_{j=2}^{N}\theta_j\varphi_j(x) : (\theta_2, \cdots, \theta_N) \in \Theta_N\}$.

Family of dim $N-1$

$\mathcal{F}_N \subset \tilde{W}(\beta, L)$ By def of Θ_N. $R_\varepsilon^* \ge \inf_{T_\varepsilon}\sup_{f \in \mathcal{F}_N}\mathbb{E}_f\|T_\varepsilon - f\|_2^2$

$\forall f \in \mathcal{F}_N$, T_ε. there exists r.v. $\hat{\theta}^N = (\hat{\theta}_2, \cdots, \hat{\theta}_N) \in \Theta_N$ s.t.

$$\|T_\varepsilon - f\|_2 \ge \|\sum_{j=2}^{N}\hat{\theta}_j\varphi_j - f\|_2. \quad \text{a.s.}$$ $\sum_{j=2}^{N}\hat{\theta}_j\varphi_j = T_\varepsilon \downarrow \mathcal{F}_N$

Let $\mathbb{E}_\theta \triangleq \mathbb{E}_{f_{\theta^N}}$.

$$R_\varepsilon^* \ge \inf_{\hat{\theta}^N \in \Theta_N}\sup_{\theta^N \in \Theta_N}\mathbb{E}_\theta\|\sum_{j=2}^{N}(\hat{\theta}_j - \theta_j)\varphi_j\|_2^2 = \inf_{\hat{\theta}^N \in \Theta_N}\sup_{\theta^N \in \Theta_N}\mathbb{E}_\theta\left[\sum_{j=2}^{N}(\hat{\theta}_j - \theta_j)^2\right]$$

2. MINIMAX $\to$ Bayes.

dist $\mu(\theta^N) = \prod_{k=2}^{N} \mu_{s_k}(\theta_k)$ $\quad s_k^2 = (1-\delta) v_k^2 \quad 0<\delta<1$.

supp on $\mathbb{R}^{N-1}$

$$R_\varepsilon^* \geq \inf_{\hat\theta^N \in \Theta_N} \sum_{k=2}^{N} \int_{\Theta_N} \mathbb{E}_\theta[(\hat\theta_k - \theta_k)^2]\, \mu(\theta^N)\, d\theta^N \geq \underset{\text{main term}}{I^*} - \underset{\text{residual term}}{r^*}$$

$$I^* = \inf_{\hat\theta^N} \sum_{k=2}^{N} \int_{\mathbb{R}^{N-1}} \mathbb{E}_\theta[(\hat\theta_k - \theta_k)^2]\, \mu(\theta^N)\, d\theta^N$$

$$r^* = \sup_{\hat\theta^N \in \Theta_N} \sum_{k=2}^{N} \int_{\Theta_N^c} \mathbb{E}_\theta[(\hat\theta_k - \theta_k)^2]\, \mu(\theta^N)\, d\theta^N.$$

Now it suffices to prove.

1. lower bounds on the main terms.
2. residual terms are negligible

$$I^* \geq C^* \varepsilon^{\frac{4\beta}{2\beta+1}} (1 + o(1)) \qquad \varepsilon \to 0.$$

$$r^* = o\left(\varepsilon^{\frac{4\beta}{2\beta+1}}\right) \qquad \varepsilon \to 0.$$

(1). main terms of Bayes risks.

$$I^* = \inf_{\hat\theta^N} \sum_{k=2}^{N} \int_{\mathbb{R}^{N-1}} \mathbb{E}_\theta[(\hat\theta_k - \theta_k)^2]\, \mu(\theta^N)\, d\theta^N$$

$$\geq \sum_{k=2}^{N} \inf_{\hat\theta_k} \int_{\mathbb{R}^{N-1}} \mathbb{E}_\theta[(\hat\theta_k - \theta_k)^2]\, \mu(\theta^N)\, d\theta^N$$

Let $\mathbb{P}_\theta = \mathbb{P}_{f_{\theta^N}}$. $\mathbb{P}_f$ be dist of $X = \{Y(t), t \in [0,1]\}$.

$\mathbb{P}_0$ is dist of Wigner Pro

Girsanov's theorem

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}(X) = \exp\left(\varepsilon^{-2} \sum_{j=2}^{N} \theta_j y_j - \frac{\varepsilon^{-2}}{2} \sum_{j=2}^{N} \theta_j^2\right) \triangleq S(y_2, \cdots, y_N, \theta^N)$$

$$y_j = \int_0^1 \varphi_j(t)\, dY(t), \quad j = 2, \cdots, N.$$