Lecture 4

$X_1, X_2, \ldots, X_n \sim \eta$ on $[0,1]$. Let $0<\alpha<1$.

$$\mathcal{P}(\alpha, M) = \left\{ \eta : \int \eta \, d\mu = 1,\ \eta \geq 0,\ |\eta(0)| \leq M,\ \sup_{x,y \in [0,1]} |\eta(x) - \eta(y)| / |x-y|^\alpha \leq M \right\}$$

→ Intersection of the set of all densities with

$$H(\alpha, M) = \left\{ f : [0,1] \to \mathbb{R} : \sup_{y,x \in [0,1]} \frac{|f(x) - f(y)|}{|x-y|} \leq M \right\}.$$

We will come to $\alpha \geq 1$ later. Let α, M be given to us. We shall also assume $\|f\|_\infty \leq M$.

Let $\{e_1, e_2, \ldots\}$ be an o.n.b. of $L_2(\mu) = \{f : \int f^2 d\mu < \infty\}$

We will make the choice of basis clear later.

For a given $k \in \mathbb{N}$, let $\Pi(x_1, x_2) = \sum_{l=1}^{k} e_l(x_1) e_l(x_2)$ be the orthogonal projection kernel onto $\text{lin}\{e_1, \ldots, e_k\}$

Let L be a given closed subspace of $L_2(\mu)$. Let

$$\Pi_L g = \underset{l \in L}{\text{argmin}} \int (g - l)^2 d\mu$$

$$\Leftrightarrow \quad \int (g - \Pi_L g) \, l \, d\mu = 0 \quad \forall \, l \in L$$

Indeed if $L = \text{lin}\{e_1, \ldots, e_k\}$ then

$$\Pi_L g(x_1) = \int \Pi(x_1, x_2) g(x_2) \, d\mu(x_2)$$

$$\Rightarrow \quad g(x_1) = \int \Pi(x_1, x_2) g(x_2) \, d\mu(x_2) \quad \forall g \in L.$$

We argued last time that we will use the "degenerate part" of $2\Pi(x_1, x_2)$ as our candidate for partial representation of $\chi_\eta^{(2)}(x_1, x_2)$. Call the degenerate $D_\eta 2\Pi(x_1, x_2)$

Degeneracy mean ~~$\mathbb{E}_\eta \Pi(x_1, x_2)$~~

$$\mathbb{E}_\eta \, D_\eta 2\Pi(X_1, x_2) = \mathbb{E}_\eta \, D_\eta 2\Pi(x_1, X_2) = 0 \quad \forall x_1, x_2$$

<u>Check</u>: $D_\eta 2\Pi(X_1, X_2) = 2\Big(\Pi(X_1, X_2) - \mathbb{E}_\eta(\Pi(X_1, X_2) \mid X_1)$

$$- \mathbb{E}_\eta(\Pi(X_1, X_2) \mid X_2) + \mathbb{E}_\eta(\Pi(X_1, X_2))\Big)$$

Call this $\Pi_\eta(X_1, X_2)$

Resulting estimator,

$$\hat{\chi}_n = \chi(\hat{\eta}) + \frac{2}{n} \sum_{i=1}^{n} \left(\hat{\eta}(X_i) - \chi(\hat{\eta})\right)$$

$$+ \frac{1}{n(n-1)} \sum_{i \neq j} \Pi_{\hat{\eta}}(X_i, X_j)$$

Let's pay some attention to $\Pi_{\hat{\eta}}(X_i, X_j)$

$$\mathbb{E}_{\hat{\eta}}(\Pi(X_1, X_2) \mid X_1) = \int \Pi(X_1, x_2) \, \hat{\eta}(x_2) \, d\mu(x_2)$$

$$= \sum_{l=1}^{k} \int e_l(x_2) \hat{\eta}(x_2) \, d\mu(x_2) \; e_l(X_2)$$

$$= \Pi_L \hat{\eta}(X_1)$$

$$\therefore \; \mathbb{E}_{\hat{\eta}}(\Pi(X_1, X_2) \mid X_2) = \Pi_L \hat{\eta}(X_2)$$

Let $\hat{\eta}(x) = \frac{1}{n}\sum_{l=1}^{k}\sum_{i=1}^{n} e_l(x)\, e_l(x_i)$

$$= \sum_{l=1}^{k}\left(\frac{1}{n}\sum_{i=1}^{n} e_l(x_i)\right) e_l(x) \in L$$

$\therefore\ \Pi_L \hat{\eta}(x_i) = \hat{\eta}(x_i)$

$\therefore\ \mathbb{E}_{\hat{\eta}}\left(\Pi(x_i, x_j) \mid x_i\right) = \hat{\eta}(x_i)$

Also, $\mathbb{E}_{\hat{\eta}}\left(\Pi(x_i, x_j)\right) = \sum_{l=1}^{k} \langle \hat{\eta}, e_l \rangle^2$

$$= \|\Pi_L \hat{\eta}\|_2^2$$

$$\chi(\hat{\eta}) = \|\hat{\eta}\|_2^2 = \|\Pi_L \hat{\eta}\|_2^2$$

$$\therefore\ \hat{\chi}_n = \chi(\hat{\eta}) + \frac{2}{n}\sum_{i=1}^{n}\left(\hat{\eta}(x_i) - \chi(\hat{\eta})\right)$$

$$+ \frac{1}{n(n-1)}\sum_{i \neq j}\Big\{\Pi(x_i, x_j) - \mathbb{E}_{\hat{\eta}}\left(\Pi(x_i, x_j) \mid x_i\right) - \mathbb{E}_{\hat{\eta}}\left(\Pi(x_i, x_j) \mid x_j\right)$$

$$+ \mathbb{E}_{\hat{\eta}}\left(\Pi(x_i, x_j)\right)\Big\}$$

$$= \chi(\hat{\eta}) - 2\chi(\hat{\eta}) + \frac{2}{n}\sum_{i=1}^{n}\hat{\eta}(x_i)$$

$$+ \frac{1}{n(n-1)}\sum_{i \neq j}\Pi(x_i, x_j) - \frac{1}{n(n-1)}\sum_{i \neq j}\hat{\eta}(x_i) - \frac{1}{n(n-1)}\sum_{i \neq j}\hat{\eta}(x_j)$$

$$+ \chi(\hat{\eta})$$

$$= \frac{1}{n(n-1)}\sum_{i \neq j}\Pi(x_i, x_j)$$

Analysis of the estimator,

$$\hat{x}_n = \frac{1}{n(n-1)} \sum_{i \neq j} \cancel{\cdot}\, \Pi(x_i, x_j)$$

$$\Pi(x_1, x_2) = \sum_{\ell=1}^{k} e_\ell(x_1)\, e_\ell(x_2)$$ over our model

$\mathcal{P}(\alpha, M)$. Indeed we want to evaluate

$$\sup_{\eta \in \mathcal{P}(\alpha, M)} \mathbb{E}_\eta \left(\hat{x}_n - x(\eta) \right)^2$$

Fix $\eta \in \mathcal{P}(\alpha, M)$. Enough to uniformly bound

$$\mathbb{E}_\eta \left(\hat{x}_n - x(\eta) \right)^2 = \mathbb{E}_\eta \left(\hat{x}_n - \mathbb{E}_\eta \hat{x}_n \right)^2 + \left(\mathbb{E}_\eta \hat{x}_n - x(\eta) \right)^2$$

$$= \mathrm{Var}_\eta\, \hat{x}_n + \left(\mathrm{Bias}_\eta\, \hat{x}_n \right)^2$$

Let's focus on the bias first.

$$\left(\mathrm{Bias}_\eta\, \hat{x}_n \right)^2 = \left(\mathbb{E}_\eta (\hat{x}_n) - x(\eta) \right)^2$$

$$= \left(\|\eta\|_2^2 - \cancel{\#}\, \|\Pi_L \eta\|_2^2 \right)^2$$

$$= \|\Pi_{L^\perp} \eta\|_2^4$$

~~Variance~~ $\mathrm{Var}_\eta\, \hat{x}_n$ is a little trickier. For that we ~~will~~ ~~#~~ make a small digression.

Write,

$$\frac{1}{n(n-1)}\sum_{i\neq j}\left(\Pi(x_i,x_j) - \mathbb{E}_\eta(\Pi(x_i,x_j))\right)$$

$$= \frac{2}{n}\sum_{i=1}^{n}\left(\mathbb{E}_\eta(\Pi(x_i,x_j)|x_i) - \mathbb{E}_\eta(\Pi(x_i,x_j))\right)$$

$$+ \frac{1}{n(n-1)}\sum_{i\neq j}\Big[\Pi(x_i,x_j) - \mathbb{E}_\eta(\Pi(x_i,x_j)|x_i) - \mathbb{E}_\eta(\Pi(x_i,x_j)|x_j) + \mathbb{E}_\eta(\Pi(x_i,x_j))\Big]$$

$$= T_1 + T_2 \qquad \mathbb{E}_\eta(T_1T_2) = 0$$

$$\therefore \operatorname{Var}_\eta \hat{x}_n = \mathbb{E}(T_1^2) + \mathbb{E}(T_2^2)$$

$$= \frac{4}{n}\mathbb{E}_\eta\left[\mathbb{E}_\eta(\Pi(x_1,x_2)|x_1) - \mathbb{E}_\eta(\Pi(x_1,x_2))\right]^2$$

$$+ \mathbb{E}\left[\frac{1}{n(n-1)}\sum_{i\neq j}\left(\Pi(x_i,x_j) - \mathbb{E}_\eta(\Pi(x_i,x_j)|x_i) - \mathbb{E}_\eta(\Pi(x_i,x_j)|x_j) + \mathbb{E}_\eta(\Pi(x_i,x_j))\right)\right]^2 \quad (*)$$

~~Let~~ For controlling the second term above we will have the following discussion.

Suppose $X_1,\dots,X_n \sim P$ on $(\Omega,\mathcal{A})$ and $f:\Omega^2 \to \mathbb{R}$ be a measurable map which is symmetric in its ~~2~~ arguments. Suppose f is also degenerate w.r.t. P i.e.

$$\int f(x_1,x_2)\,dP(x_1) = \int f(x_1,x_2)\,dP(x_2) = 0 \quad \text{for all } x_1,x_2$$

then

$$P^n(U_nf)^2 = \frac{\cancel{\frac{1}{n(n-1)^2}}\ \cancel{\frac{n}{2}}\ P^2f^2}{\binom{n}{2}}$$

Proof: $P^n(U_n f)^2 = \dfrac{1}{n^2(n-1)^2} \sum_{\substack{(i_1,j_1),\\(i_2,j_2)}} \mathbb{E}\left(f(x_{i_1},x_{j_1})\, f(x_{i_2},x_{j_2})\right)$

Now consider the case where ~~there~~ at least one of the indices is not repeated. Any such term vanishes by independence & degeneracy. The only terms that survive are when $(i_1,j_1) = (i_2,j_2)$. The number of such cases $= n(n-1)$ which gives the result by symmetry. (note that both $((i,j), (j,i))$ & $((j,i),(i,j))$ contribute) ∎

For us, $f(x_1,x_2) = \Pi(x_1,x_2) - \mathbb{E}_\eta\left(\Pi(x_1,x_2) \mid x_1\right)$
$$- \mathbb{E}_\eta\left(\Pi(x_1,x_2) \mid x_2\right)$$
$$+ \mathbb{E}_\eta\left(\Pi(x_1,x_2)\right)$$

Now $\mathbb{E}_\eta\left(\mathbb{E}_\eta\left(\Pi(x_1,x_2)\mid x_1\right)\right)^2 \le \mathbb{E}\left(\Pi^2(x_1,x_2)\right)$ by Jensen's Inequality.

Also, $(a+b+c+d)^2 \le 2(a+b)^2 + 2(c+d)^2$
$$\le 4(a^2+b^2+c^2+d^2)$$

$\Rightarrow \mathbb{E}_\eta\left(f^2(x_1,x_2)\right) \le 16\, \mathbb{E}_\eta\left(\Pi^2(x_1,x_2)\right)$

Actually $f(x_1,x_2)$ is the projection of $\Pi(x_1,x_2)$ onto to space of degenerate functions of 2 variables & projections contract norm. But the above calculation is sufficient for our purposes.

Going back to (*),

$$\mathrm{Var}_\eta \hat{X}_n \le \frac{4}{n}\, \mathbb{E}_\eta\Big[\mathbb{E}_\eta\big(\Pi(x_1,x_2)\,|\,x_1\big) - \mathbb{E}_\eta\big(\Pi(x_1,x)\big)\Big]^2 + \frac{32}{n(n-1)}\, \mathbb{E}_\eta\big(\Pi^2(x_1,x_2)\big)$$

~~Now,~~ $\cancel{\mathbb{E}_\eta \big[\mathbb{E}_\eta}$

$$\le \frac{4}{n}\, \mathbb{E}_\eta\Big[\mathbb{E}_\eta^2\big[\Pi(x_1,x_2)\,|\,x_1\big]\Big] + \frac{64}{n^2}\, \mathbb{E}_\eta\big[\Pi^2(x_1,x_2)\big]$$

Now, $\mathbb{E}_\eta\big[\Pi(x_1,x_2)\,|\,x_1\big] = \Pi_L\, \eta(x_1)$

$$\Rightarrow \mathrm{Var}_\eta \hat{X}_n \le \frac{4}{n}\, \mathbb{E}_\eta\Big(\big(\Pi_L\, \eta(x_1)\big)^2\Big) + \frac{64}{n^2}\, \mathbb{E}_\eta\big(\Pi^2(x_1,x_2)\big)$$

Now, $|\eta(x) - \eta(y)| \le M\,|x-y|^\alpha \quad \forall\, x,y \in [0,1]$ & $|\eta(0)| \le M$

$$\Rightarrow \|\eta\|_\infty \le 2M$$

$$\therefore \mathbb{E}_\eta\Big(\big(\Pi_L\, \eta(x_1)\big)^2\Big) = \int \big(\Pi_L \eta(x_1)\big)^2\, \eta(x_1)\, d\mu(x_1)$$

$$\le 2M\, \|\Pi_L \eta\|_2^2 \le 2M\, \|\eta\|_2^2 \le 8M^3$$

$$\Rightarrow \mathrm{Var}_\eta \hat{X}_n \le \frac{32M^3}{n} + \frac{64}{n^2}\, \mathbb{E}_\eta\big(\Pi^2(x_1,x_2)\big)$$

Now, $\mathbb{E}_\mu\left(\Pi^2(x_1,x_2)\right)$

$$= \iint\left(\sum_{\ell=1}^{k} e_\ell(x_1)\, e_\ell(x_2)\right)^2 d\mu(x_1)\, d\mu(x_2)$$

$$= \iint \sum_{\ell,\ell'} e_\ell(x_1)\, e_{\ell'}(x_1)\, e_\ell(x_2)\, e_{\ell'}(x_2)\, d\mu(x_1)\, d\mu(x_2)$$

$$= \sum_{\ell=1}^{k} \|e_\ell\|_2^2\, \|e_{\ell'}\|_2^2 = k$$

We need $\mathbb{E}_\eta\left(\Pi^2(x_1,x_2)\right)$ when $P_\eta \ll \mu$ with $\|\eta\|_\infty = \left\|\frac{dP_\eta}{d\mu}\right\|_\infty \le M$.

We make use of the following lemma.

Lemma: $\mathbb{E}_\eta\left(\Pi^2(x_1,x_2)\right) \le 4\left(1+\|\eta\|_\infty\right)^4 \mathbb{E}_\mu\left(\Pi^2(x_1,x_2)\right)$

Actually for any measurable $f: \Omega^m \to \mathbb{R}$ & probability measures $P \ll Q$,

$$P^n (U_n f)^2 \le 2m\left(1+\left\|\frac{dP}{dQ}\right\|_\infty\right)^{2m} Q^n (U_n f)^2$$

We will prove the lemma later. Now let us apply the lemma to analyze our variance.

$$\mathrm{Var}_\eta \hat{\chi}_n \le \frac{324 M^3}{n} + \frac{64}{n^2}\, \mathbb{E}_\eta\left(\Pi^2(x_1,x_2)\right)$$

$$\le \frac{324 M^3}{n} + \frac{256}{n^2}(1+M)^4\, \mathbb{E}_\mu\left(\Pi^2(x_1,x_2)\right)$$

$$= \frac{324 M^3}{n} + \frac{256}{n^2}(1+M)^4 k$$

$$\le C(M)\left(\frac{1}{n} + \frac{k}{n^2}\right)$$

$\Rightarrow$ We have

$$\left(\text{Bias}_\eta \hat{x}_n\right)^2 \le \|\Pi_{L^\perp} \eta\|_2^4$$

$$\text{Var}_\eta \hat{x}_n \le C(M)\left(\frac{1}{n} + \frac{k}{n^2}\right)$$

Now, $\exists$ basis of $L_2(\mu)$ s.t. $f \in H(\alpha, M)$

$$\Rightarrow \|\Pi_{L^\perp} f\|_\infty \le \underset{\substack{\uparrow \\ \text{universal constant}}}{c^*} M k^{-\beta}$$

(We will talk about such ~~spaces~~ bases soon)

For such bases, $\eta \in \mathcal{P}(\alpha, M)$

$\cancel{\Rightarrow \text{Bias}_\eta \hat{x}_n}$

$$\Rightarrow \mathbb{E}_\eta\left(\hat{x}_n - x(\eta)\right)^2 \le C^* M k^{-\beta} + C(M)\left(\frac{1}{n} + \frac{k}{n^2}\right)$$

$\boxed{\text{So}}$, if $\alpha \le \frac{1}{4}$ choose $k = n^{2/4\beta+1}$ $\Rightarrow$ MSE $\le C'(M)\, n^{-8\beta/4\beta+1}$

if $\alpha > 1/4$ choose $k = n$ $\Rightarrow$ MSE $= C'(M)\frac{1}{n}$

$$\mathbb{E}_P\left(f^2(x_1,x_2)\right)$$

$$= E_Q\left(f^2(x_1,x_2)\frac{dP}{dQ}(x_1)\frac{dP}{dQ}(x_2)\right)$$

$$= E_Q\left(f^2(x_1,x_2)\,\Lambda(x_1,x_2)\right) \qquad \Lambda(x_1,x_2) = \frac{dP}{dQ}(x_1)\frac{dP}{dQ}(x_2)$$

$$= E_Q\left(f^2(x_1,x_2)\left(\Lambda(x_1,x_2) - \mathbb{E}_Q(\Lambda(x_1,x_2)) + \mathbb{E}_Q(\Lambda(x_1,x_2))\right)\right)$$

$$\begin{aligned}= E_Q\Big(f^2(x_1,x_2)\Big(&\mathbb{E}_Q(\Lambda(x_1,x_2)\mid x_1) + \mathbb{E}_Q(\Lambda(x_1,x_2)\mid x_2) \\ &- 2\mathbb{E}_Q(\Lambda(x_1,x_2)) \\ &+ \Lambda(x_1,x_2) - \mathbb{E}_Q(\Lambda(x_1,x_2)\mid x_2) - \mathbb{E}_Q(\Lambda(x_1,x_2)\mid x_2) \\ &+ \mathbb{E}_Q(\Lambda(x_1,x_2)) + \mathbb{E}_Q(\Lambda(x_1,x_2))\Big)\Big)\end{aligned}$$

$$\begin{aligned}= \mathbb{E}_Q(f^2(x_1,x_2)) + \mathbb{E}_Q\Big(f^2(x_1,x_2)\,\big(&\Lambda(x_1,x_2) - \mathbb{E}_Q(\Lambda\mid x_1) \\ &- \mathbb{E}_Q(\Lambda\mid x_2) + \mathbb{E}_Q(\Lambda)\big)\Big)\end{aligned}$$

$$\begin{aligned}\le E_Q(f^2(x_1,x_2)) &+ \mathbb{E}_Q\, f^2(x_1,x_2)\,\|\Lambda(x_1,x_2)\|_\infty \\ &+ 2\mathbb{E}_Q\, f^2(x_1,x_2)\,\|\cancel{\Lambda(x_1,x_2)}\; E_Q(\Lambda\mid x_1)\|_\infty \\ &+ \mathbb{E}_Q(f^2(x_1,x_2))\end{aligned}$$

$$\begin{aligned}= 2\,\mathbb{E}_Q(f^2(x_1,x_2)) &+ \mathbb{E}_Q\left(f^2(x_1,x_2)\left\|\frac{dP}{dQ}\right\|_\infty^2\right) \\ &+ 2\,\mathbb{E}_Q\left(f^2(x_1,x_2)\left\|\frac{dP}{dQ}\right\|_\infty\right)\end{aligned}$$

$$\le 2\left(1+\left\|\frac{dP}{dQ}\right\|_\infty\right)^2 E_Q\left(f^2(x_1,x_2)\right)$$