Lecture 15

In this lecture we finish construction of adaptive honest confidence sets in L_2 over Besov balls.

Fix a CDV family of wavelet basis with regularity $r \geq \lceil \beta_{max} \rceil$ and define

$$B^{\beta}_{2\infty}(M) := \left\{ f \in L_2[0,1] : \sup_{l \geq J_0} 2^{l\beta} \left(\| \langle f, \psi_{l\cdot} \rangle \|_2 \right) \leq M \right\}$$

We want to construct honest adaptive confidence sets in L_2-norm over $\bigcup_{\beta \in [\beta_{min}, \beta_{max}]} \mathcal{P}(\beta, M, B)$ where

$$\mathcal{P}(\beta, M, B) = \left\{ f \in B^{\beta}_{2\infty}(M) : 0 \leq f \leq B \text{ a.s. } \lambda, \int f = 1 \right\}$$

Mathematically, if we have a sample of size n, $X_1, \dots, X_n$ having density η w.r.t λ on $[0,1]$ then we want to construct $\hat{C}_n \subseteq L_2[0,1]$ such that the following hold.

(i) $\liminf_{n \to \infty} \inf_{\eta \in \mathcal{P}(\beta_{min}, M, B)} \mathbb{P}_\eta (\eta \in \hat{C}_n) \geq 1 - \alpha$

(ii) $\limsup_{n \to \infty} \sup_{\eta \in \mathcal{P}(\beta, M, B)} \mathbb{P}_\eta \left(\frac{\text{diam}(\hat{C}_n)}{n^{-\beta/2\beta+1}} > c(M, B) \right) \leq \alpha$ for all $\beta \in [\beta_{min}, \beta_{max}]$

According to lecture 14, this can not be achieved if $\beta_{max} \geq 2\beta_{min}$ (from our testing argument). What is the best we can do in this case?

(Note: $n^{-\beta/2\beta+1}$ is the asymptotic minimax rate of estimation of η over $B^{\beta}_{2\infty}(M)$.)

We now define parameter spaces over which we will produce honest adaptive confidence sets in L_2 when $\beta_{max} \geq 2\beta_{min}$. The method is best explained assuming $\beta_{max} = 2^N \beta_{min}$ for some $N > 1$. Let $\beta_j = 2^{j-1}\beta_{min}$ $j = 1, \dots, N+1$ and let $\rho_n(\beta_j) = C^* n^{-\frac{2\beta_j}{4\beta_j+1}}$ $j = 1, \dots, N+1$ where C^* is large enough (depending on M, B, α so that the testing problem (with $\beta' > \beta_j$) below has error less than $\alpha/4N$ for a suitable test.

$$H_0: \eta \in \mathcal{P}(M, B, \beta') \quad vs \quad H_1(\rho_n): \eta \in \mathcal{P}(M, B, \beta_j), \inf_{\eta' \in H_0} \|\eta - \eta'\| \geq \rho_n(\beta_j)$$

This can be done according to our test construction last time.

Let $\mathcal{F}_N^{(n)} = B_{2\infty}^{\beta}(\beta_{max}) \cup \bigcup_{j=1}^{N+1} \bar{B}_{2\infty}^{\beta_{j-1}}(M, \rho_n(\beta_{j-1}))$ where

$$\bar{B}_{2\infty}^{\beta_{j-1}}(M, \rho_n(\beta_{j-1})) = \left\{ h \in B_{2\infty}^{\beta_{j-1}}(M) : \| h - B_{2\infty}^{\beta_l}(M) \| \geq \rho_n(\beta_{j-1}) \ \forall\ l = j, \dots, N+1 \right\}$$

Note that trivially $\mathcal{F}_N^{(n)} \subseteq \bigcup_{\beta \in [\beta_{min}, \beta_{max}]} \mathcal{P}(\beta, M, B)$. Also N does depend on n. Therefore we can show that as $n \to \infty$ $\mathcal{F}_N^{(n)}$ grows dense in $\bigcup_{\beta \in [\beta_{min}, \beta_{max}]} \mathcal{P}(\beta, M, B)$, which is the full model.

An important question is whether $\mathcal{F}_N^{(n)}$ was taken to grow as fast as possible with n. In other words, whether a smaller choice of $\rho_n(\beta_j)$ would have been possible. If $\hat{C}_n$ is an honest confidence set over $\mathcal{F}_N^{(n)}$, which has a faster separation rate $\rho_n'(\beta_j) = o(\rho_n(\beta_j))$ for some j, then we can $\hat{C}_n$ to test $H_0: \eta \in \mathcal{P}(\beta, M, B)$ vs

$H_1(\rho_n'(\beta_j))$ for some $\beta > 2\beta_j$, which is a contradiction.
Henceforth we will therefore consider honest adaptation over $\mathcal{F}_N^{(n)}$. In particular, we will prove the following theorem.

<u>Theorem</u>: Let $\beta_{max} > 2\beta_{min}$ be such that $\exists\ N > 1$ with $\beta_{max} = 2^N \beta_{min}$. Then $\exists$ a confidence set based on $X_1, \ldots, X_n, \beta_{min}, \beta_{max}, M, B, \alpha$ such that the following hold.

(i) $\liminf_{n\to\infty} \inf_{\eta \in \mathcal{F}_N^{(n)}} \mathbb{P}_\eta(\eta \in \hat{C}_n) \geq 1 - \eta$

(ii) $\overline{\lim}_{n\to\infty} \sup_{\eta \in \mathcal{P}(\beta, M, B) \cap \mathcal{F}_N^{(n)}} \mathbb{P}_\eta\left(\text{diam}(\hat{C}_n) > c^* n^{-\beta/2\beta+1}\right) \leq \alpha$ for all $\beta \in [\beta_{min}, \beta_{max}]$

where $c^* = c(\alpha, M, B, \beta_{min}, \beta_{max})$.

<u>Proof</u>: Let $\mathcal{F}_j^{(n)} = \bar{B}_{2\infty}^{\beta_{j-1}}(M, \rho_n(\beta_{j-1}))$ $j = 1, \ldots, N+1$
Recall the testing problem, for $\beta > \beta'$

$H_0: \eta \in B_{2\infty}^{\beta}(M)$ vs $H_1: \eta \in B_{2\infty}^{\beta'}(M)$ s.t. $\inf_{\eta' \in H_0} \|\eta - \eta'\|_2 \geq \rho_n(\beta')$

Call the optimal test $\Psi_n(\beta, \beta')$ at error $\alpha/4N$.
We first test $H_0: \eta \in \mathcal{F}_2^{(n)}$ vs $H_1: \eta \in \mathcal{F}_1^{(n)}$ at error ~~rate~~ $\alpha/4N$

If we reject we set $\hat{\beta} = \beta_1 = \beta_{min}$ and stop. Otherwise continue, and at the jth step test

$H_0: \eta \in \mathcal{F}_{j+1}^{(n)}$ vs $H_1: \eta \in \mathcal{F}_j^{(n)}$

using ~~$\Psi_n(\beta_j, \beta_{j-1})$~~ $\Psi_n(\beta_j, \beta_{j-1})$. If we reject, stop and declare $\hat{\beta} = \beta_{j-1}$.

If non of the hypotheses are rejected we declare $\hat{\beta} = \beta_N$.
This procedure determines the shell η belongs to.
Once we have this location, we construct the confidence set using our idea of risk estimation.

Without loss of generality assume we have data $X_1, \dots, X_{2n}$ and we construct an estimat $\hat{\eta}$ of η from $X_{n+1}, \dots, X_{2n}$ satisfying

$$\limsup_{n\to\infty} \sup_{\eta \in \mathcal{P}(\beta, M, B)} \mathbb{E}_\eta (\hat{\eta} - \eta)^2 \leq C(B, M)\, n^{-\frac{2\beta}{4\beta+1}} \quad \text{for all } \beta \in [\beta_{min}, \beta_{max}]$$

$$\text{and} \quad \liminf_{n \to \infty} \sup_{\eta \in \mathcal{P}(\beta, M, B)} \mathbb{P}_\eta \left(\hat{\eta} \notin B_{2\infty}^{\beta}(C(M, B)) \right) \leq \frac{\alpha}{8N}$$

$$\text{and} \quad \limsup_{n \to \infty} \sup_{\eta \in \mathcal{P}(\beta, M, B)} \mathbb{P}_\eta \left(\|\hat{\eta}\|_\infty > 2B \right) \leq \alpha/8N .$$

Refer to Bull & Nickl (2013) for such a construction. (You can also check Mukherjee & Sen (2016) for further details).

Using the first half of sample construct $\hat{\beta}$ as described above. Now for $\eta \in \mathcal{F}_N^{(n)}$ let $i_0(\eta)$ denote the unique index such that $\eta \in \mathcal{F}_{i_0+1}^{(n)}$. We first show that uniformly over $\eta \in \mathcal{P}(\beta_{min}, M, B) \cap \mathcal{F}_N^{(n)}$ $\mathbb{P}_P(\hat{\beta} \neq \beta_{i_0})$ is well controlled. Indeed if $\hat{\beta} < \beta_{i_0}$ implies that atleast one of the tests considered above are rejected. the true null hypothesis for $j = 1, \dots, i_0 - 1$.

Therefore
$$\limsup_{n\to\infty} \sup_{\eta \in \mathcal{H}^{(n)}_{i_0+1}} \mathbb{P}_\eta(\hat{\beta} < \beta_{i_0})$$
$$\le \limsup_{n\to\infty} \sum_{i<i_0} \sup_{\eta\in \mathcal{H}^{(n)}_{i_0+1}} \mathbb{P}_\eta(\hat{\beta} = \beta_i)$$
$$\le \alpha/4$$

Similarly $\limsup_{n\to\infty} \sup_{\eta\in\mathcal{H}^{(n)}_{i_0+1}} \mathbb{P}_\eta(\hat{\beta} > \beta_{i_0}) \le \alpha/4$. (looking the power of the tests)

$$\Rightarrow \limsup_{n\to\infty} \sup_{\eta\in\mathcal{H}^{(n)}_{i_0+1}} \mathbb{P}_\eta(\hat{\beta} \ne \beta_{i_0}) \le \alpha/2 .$$

~~⇒ ...~~ Now you set $\hat{j} = \left\lceil \frac{2}{4\hat{\beta}+1}\log_2 n \right\rceil$ i.e. $2^{\hat{j}} \sim n^{\frac{2}{4\hat{\beta}+1}}$.

~~Let $\hat{U}_n = \frac{1}{n(n-1)} \sum_{1\le i_1 \ne i_2 \le n} K_{V_j}(X_{i_1}, X_{i_2})$~~

Now let,
$$\hat{U}_n = \frac{1}{n(n-1)} \sum_{1\le i_1\ne i_2\le n} \sum_{l=J_0}^{\hat{j}} \sum_{k\in\mathcal{Z}_l} \left(\psi_{lk}(X_{i_1}) - \langle \psi_{lk}, \hat{\eta}\rangle\right)\left(\psi_{lk}(X_{i_2}) - \langle\psi_{lk}, \hat{\eta}\rangle\right)$$
→ unbiased estimator of $\|\Pi_{V_{\hat{j}}}(\eta - \hat{\eta})\|_2^2$ given $\hat{\eta}$. Let

$$C_n(\hat{\beta}) = \left\{\eta : \|\eta - \hat{\eta}\|_2^2 \le \hat{U}_n + C(M,B)\, n^{-\frac{4\hat{\beta}}{4\hat{\beta}+1}} + z_\alpha \tilde{\tau}_n(\eta)\right\}$$

$$\tilde{\tau}_n(\eta) = \frac{C_1}{n}\|\eta - \hat{\eta}\|_2^2 + C_2 \frac{2^{\hat{j}}}{n(n-1)}$$ for suitable constants C, C_1, and C_2

Now, $\mathbb{P}_\eta(\eta \in C_n(\hat{\beta})) \ge \mathbb{P}_\eta(\eta \in C_n(\beta_{i_0})) - \frac{\alpha}{2}$

$\Rightarrow$ It's enough to show $\mathbb{P}_\eta(\eta \in \hat{C}_n(\beta_{i_0})) \geq 1 - \alpha/2$

First,

$$\mathbb{E}_{\eta,2}(\hat{U}_n) = \| \Pi_{V_j}(\eta - \hat{\eta}) \|_2^2$$

where the subscript 2 means we are conditioning on the $X_{n+1}, \ldots, X_{2n}$.

standard variance calculation shows,

$$\mathrm{Var}_{\eta,2}(\hat{U}_n) \leq \frac{c_1}{n} \| \Pi_{V_j}(\eta - \hat{\eta}) \|_2^2 + c_2 \frac{2^j}{n(n-1)}$$

where c_1 and c_2 are determined based on n, (φ, ψ) and B.

$$\leq \tilde{\tau}_n(\eta)$$

$$\Rightarrow \mathbb{P}_{\eta,2}\left(|\hat{U}_n - \mathbb{E}_{\eta,2}(\hat{U}_n)| > z_\alpha \tau_n(\eta) \right) \leq \frac{1}{z_\alpha^2}$$

Let $z_\alpha^2 = 2/\alpha$.

Also, $\mathbb{E}_{\eta,2}(\hat{U}_n) = \| \Pi_{V_j}(\eta - \hat{\eta}) \|_2^2$

$$= \| \eta - \hat{\eta} \|_2^2 - \| \Pi_{V_j^\perp}(\eta - \hat{\eta}) \|_2^2$$

Now $\hat{\eta} \in B_{2\infty}^{\beta_{i_0}}(C(M, B))$ with high probability.

For $f \in B_{2\infty}^{\beta}(M)$, $\| \Pi_{V_j^\perp}(\eta) \|_2^2 = \sum_{l=j}^{\infty} \| \Pi_{W_l}(\eta) \|_2^2$

$$\leq M/2^{-2l\beta} \leq C(M)\, 2^{-2j\beta}$$

$$\Rightarrow \mathbb{E}_{\eta,2}(\hat{U}_n) \geq \| \eta - \hat{\eta} \|_2^2 - C(M)\, 2^{-2j\beta_{i_0}}$$

$$= \| \eta - \hat{\eta} \|_2^2 - C(M)\, n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}}$$

because we are on the set where $\hat{\beta} = \beta_{i_0}$.

$$\Rightarrow \mathbb{P}_{\eta,2}\left(\|\eta - \hat{\eta}\|_2^2 \le \hat{U}_n + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + z_\alpha r_n(\eta) \right) \ge 1 - \alpha/2$$

as required. It remains to show that the diameter of $\hat{C}_n$ is well controlled. Indeed the true $\beta \in [\beta_{i_0}, \beta_{i_0+1})$. The deterministic term in the diameter is $n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} = O\left(n^{-\frac{\beta}{2\beta+1}}\right)$ as $\beta_{i_0+1} = 2\beta_{i_0}$. In particular, any $\eta' \in \hat{C}_n(\beta_{i_0})$ satisfies,

$$\|\eta' - \hat{\eta}\|_2^2 \le \hat{U}_n + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + z_\alpha \left(c_1 \frac{\|\eta' - \hat{\eta}\|_2^2}{n} + c_2 \frac{2^j}{n(n-1)} \right)^{\frac{1}{2}}$$

$$\le \hat{U}_n + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + z_\alpha c_1 \frac{\|\eta - \hat{\eta}\|_2}{\sqrt{n}}$$

$\Rightarrow$ ~~$\left(1 + \frac{z_\alpha c_1}{\sqrt{n}}\right)$~~ Letting $\|\eta' - \hat{\eta}\|_2 = x_n$, we have

$$x_n^2 \le \hat{U}_n + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + \frac{z_\alpha c_1}{\sqrt{n}} x_n$$

$$\Rightarrow x_n \le 2\left(\hat{U}_n + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}}\right) + 2 \frac{z_\alpha^2 c_1^3}{n}$$

$\left(x \le b + a\sqrt{x},\ x, a, b \ge 0 \Rightarrow x^2 \le 2b + 2a^2 \right).$

~~$\frac{z_\alpha c_1^3}{n} + 2\mathbb{E}_{\eta,2}(\hat{R}_n)$~~

$$\le C^* \|\eta - \hat{\eta}\|_2^2 + C(M,B) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + \frac{z_\alpha c_1}{\sqrt{n}} x_n + \left(c_1 \frac{\|\eta - \hat{\eta}\|_2}{\sqrt{n}} + c_2 \frac{2^j}{n(n-1)} \right) z_\alpha$$

$$\le C^* n^{-\frac{\beta}{2\beta+1}} + \left(C(M,B) + c_2\right) n^{-\frac{4\beta_{i_0}}{4\beta_{i_0}+1}} + \frac{z_\alpha c_1}{\sqrt{n}} x_n \qquad \text{w.p.} \ge 1 - \frac{\alpha}{2}$$

$$\Rightarrow x_n \le 2C^* \left(n^{-\frac{\beta}{2\beta+1}} + \frac{C(M,B)+c_2}{C^*} n^{-\frac{4\beta i_0}{4\beta i_0+1}} \right) + \frac{2 z_\alpha^2 \sigma^2}{n}$$

$\Rightarrow$ Any $\eta' \in \hat{C}_n(\beta i_0)$ satisfies, w.p. $\ge 1 - \alpha/2$

$\|\eta' - \hat{\eta}\|_2^2 \le C(M,B)\, n^{-\frac{\beta}{2\beta+1}}$ as $\beta \in [\beta_{i_0}, \beta_{i_0+1})$ and $\beta_{i_0+1} = 2\beta_{i_0}$.

This completes the proof of the theorem.