STA3000F Lecture 18.

$$\log N(\delta; \Sigma(\beta), \|\cdot\|_\infty) \leq \left(\frac{c}{\delta}\right)^{1/\beta}$$

$$G_n(r) \leq \inf \left\{ c\delta_0 + \frac{c}{\sqrt{n}} \int_{\delta_0}^{r} \sqrt{\log N(\delta; \Sigma(\beta), \|\cdot\|_\infty)}\, d\delta \right\}$$

$r_n^2 = G_n(r_n)$ $\quad\quad$ $\phi_n = G_n$

$\beta > \frac{1}{2}$	$\delta_0 = 0$	$r_n = n^{-\frac{\beta}{2\beta+1}}$	$\phi_n(r) = \frac{1}{\sqrt{n}} \cdot r^{1-\frac{1}{2\beta}}$
$\beta = \frac{1}{2}$	$\delta_0 = 1/n$	$r_n = (\log n)^{1/2} \cdot n^{-1/4}$	$\phi_n(r) = \frac{\log n}{\sqrt{n}}$
$\beta < \frac{1}{2}$	$\delta_0 = n^{-\beta}$	$r_n = n^{-\beta/2}$	$\phi_n(r) = n^{-\beta}$.

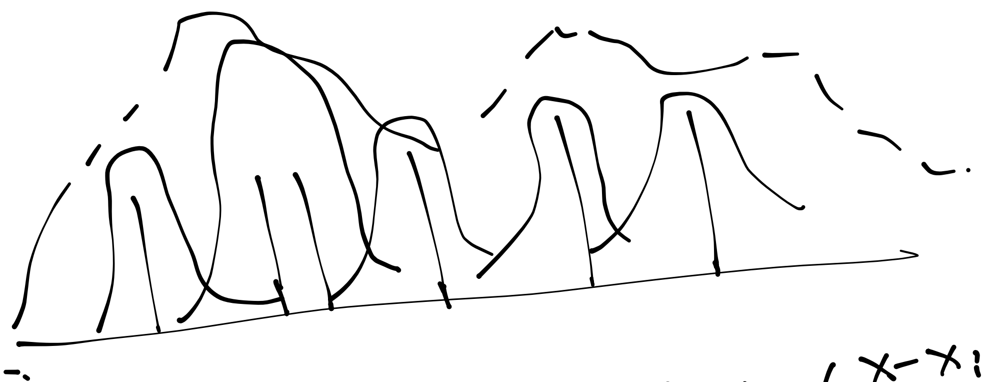
$$r_n^* = n^{-\frac{\beta}{2\beta+1}}$$

For dimension d, $\quad r_n^* = n^{-\frac{\beta}{2\beta+d}}$.

Achieved by LS when $\beta > d/2$.

Density estimation.

$$X_1, \dots, X_n \overset{iid}{\sim} P.$$



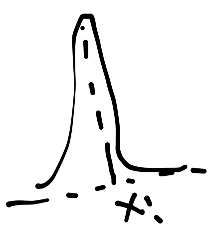
only averages within $[x_0-h, x_0+h]$

$$\hat{p}_n(x) := \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$$

where K is a kernel.

- h: Bandwidth

Small h $\qquad$ Large h



$$\int_{\mathbb{R}} \hat{p}_n(x)\,dx = \frac{1}{nh}\sum_{i=1}^{n}\int K\left(\frac{x-X_i}{h}\right)dx = \underbrace{\int K(x)\,dx = 1}_{\text{Assumption on } K}$$

Analysis of $\hat{p}_n$. x_0 fixed.

$$\mathbb{E}\left[\left|\hat{p}_n(x_0) - p(x_0)\right|^2\right] = \operatorname{var}\left(\hat{p}_n(x_0)\right) + \left|\mathbb{E}[\hat{p}_n(x_0)] - p(x_0)\right|^2.$$

$$\operatorname{var}\left(\hat{p}_n(x_0)\right) = \frac{1}{nh^2} \operatorname{var}\left(K\left(\frac{X-x_0}{h}\right)\right)$$

$$\leq \frac{1}{nh^2} \int K^2\left(\frac{y-x}{h}\right) p(y)\, dy$$

$$\leq \frac{1}{nh} \cdot p_{max} \boxed{\int_{\mathbb{R}} K^2(x)\, dx}$$ Assumption on K.

- Bias. $\left|\mathbb{E}[\hat{p}_n(x_0)] - p(x_0)\right|$

$$= \left|\frac{1}{h} \int_{\mathbb{R}} K\left(\frac{y-x_0}{h}\right) \left(p(y) - p(x_0)\right) dy\right|$$

$$\leq \int_{\mathbb{R}} |K(u)| \cdot \left|p(x_0 + uh) - p(x_0)\right| du.$$

Naive bound $p \in \Sigma(\beta)$ $(0 < \beta \leq 1)$.

$$\left|p(x_0 + uh) - p(x_0)\right| \leq h^\beta \cdot |u|^\beta$$

$$\left|\mathrm{Bias}(x_0)\right| \leq h^\beta \underbrace{\int_{\mathbb{R}} |K(x)| \cdot |u|^\beta\, du}_{< +\infty}$$ Assumption on K.

$$\text{MSE} \le \text{var} + \text{bias}^2$$

$$\le \frac{1}{nh} \cdot p_{max} \int K^2 + h^{2\beta} \left[\int |K| \cdot |u|^\beta \right]$$

$$\left(h_n^* = n^{-\frac{1}{2\beta+1}} \right)$$

$$\le C \cdot n^{-\frac{2\beta}{2\beta+1}}$$

for $\beta \in (0, 1]$.

$\beta > 1$. $\quad l = \lfloor \beta \rfloor$

$$p(x_0 + uh) - p(x_0)$$

$$= \partial p(x_0) \cdot uh + \frac{\partial^2 p(x_0)}{2!} (uh)^2 + \cdots + \frac{\partial^{l-1} p(x_0)}{(l-1)!} \cdot (uh)^{l-1}$$

$$+ \frac{\partial^l p(x_0 + \tau h u)}{l!} (uh)^l$$

for some $\tau \in [0, 1]$.

$$\text{Bias} = \int_{\mathbb{R}} K(u) \left(p(x_0 + uh) - p(x_0) \right) du$$

Assumption on K: $\quad j = 1, 2, \dots l,$

$$\int u^j \cdot K(u) \, du = 0.$$

$$\left|\mathrm{Bias}(x_0)\right| \leq \int_{\mathbb{R}} \frac{|K(u)|}{l!} \cdot \left|\partial^l p(x_0+\tau uh) - \partial^l p(x_0)\right| \cdot (uh)^l \, du$$

$$\leq \int_{\mathbb{R}} \frac{|K(u)| \cdot |u|^{\beta-l} \cdot |u|^l}{l!} \cdot h^{\beta-l+l} \, du$$

$$\leq h^\beta \cdot \underbrace{\int_{\mathbb{R}} |K(u)| \cdot |u|^\beta du}_{\text{Assume } <+\infty}$$

$$\mathrm{MSE}(x_0) \leq n^{-\frac{2\beta}{2\beta+1}} \qquad \left(h_n^* = n^{-\frac{1}{2\beta+1}}\right).$$

Construction of Kernel.

Legendre poly: orthonormal polys in $L^2([-1,1])$.

$$\varphi_0(x) = \frac{1}{\sqrt{2}} \qquad \varphi_m = \sqrt{\frac{2m+1}{2}} \frac{1}{2^m m!} \frac{d^m}{dx^m}\left((x^2-1)^m\right).$$

$$\int_{-1}^{1} \varphi_j \varphi_l = \begin{cases} 1 & j = l \\ 0 & \text{otherwise} \end{cases}$$

$$K(u) = \sum_{m=0}^{l} \varphi_m(0) \cdot \varphi_m(u) \mathbb{1}\{u \in [-1,1]\}$$

$$\int_{\mathbb{R}} K(x) = \int_{\mathbb{R}} \varphi_0(0) \cdot \varphi_0(u) \mathbb{1}_{u\in[-1,1]} du = 1$$

$$\int K^2 < +\infty, \qquad \int |K|\,|u|^{\beta} < +\infty.$$

$1 \le k \le l.$

$$\int x^k K(x)\,dx = \int_{-1}^{1} \left(\sum_{j=0}^{k} b_j \varphi_j(x)\right) \cdot \left(\sum_{m=0}^{l} \varphi_m(0) \cdot \varphi_m(x)\right) dx$$

$$= \sum_{j=0}^{k} b_j \varphi_j(0) = x^k \Big|_{x=0} = 0.$$