Lecture 22: Gaussian Lipschitz Concentration

Isabella Zhu

29 April 2025

§1 Recap and Proof

Last lecture we went over the setup

$$y_i = f^*(x_i) + \sigma w_i, \quad w_i \in N(0,1)$$

Non-parametric least squares

$$\hat{f} \in \underset{f \in F}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} (y_i - f(x_i))^2$$

We assume that $f^* \in F$ which means that it's well-specified. We care about the behavior

$$||\hat{f} - f^*||_n^2 = \frac{1}{n} \sum_{i=1}^{n} (\hat{f}(x_i) - f^*(x_i))^2$$

where $||\hat{f} - f^*||_n^2$ is the $L^2(\mathbb{P}_n)$-squared error. We will prove

Theorem 1.1

We have

$$||\hat{f} - f||_n \leq c(\delta_n^2 + t\delta_n)$$

with high probability in $t \geq 1$.

Proof. Recall the notion of localized complexity where

$$G_n(\delta) = \mathbb{E}_w \sup_{||f - f^*||_2 \leq \delta} \left| \frac{1}{n} \sum_{i=1}^{n} w_i (f(x_i) - f^*(x_i)) \right|$$

and we find the point where

$$\frac{G_n(\delta)}{\delta} = \frac{\delta}{2\sigma}$$

and call the solution to this δ_n. Notice that δ_n is fixed since we are assuming that x_i are fixed.

Claim 1.2 — $G_n(\delta)/\delta$ is a decreasing function when f is star-shaped.

The proof of the main claim is that

$$||\hat{f} - f^*||_n^2 \leq 2\sigma \frac{1}{n} \sum_{i=1}^{n} w_i(\hat{f}(x_i) - f^*(x_i))$$

If $||\hat{f} - f^*||_n \leq \delta_n$, then we are done. Then, we know

$$0 < \frac{\delta_n}{||\hat{f} - f^*||_n} < 1$$

We have that

$$\begin{aligned}
\frac{2\sigma}{n} \sum_{i=1}^{n} w_i(\hat{f}(x_i) - f^*(x_i)) &= 2\sigma \frac{||\hat{f} - f^*||_n}{\delta_n} \frac{1}{n} \sum_{i=1}^{n} w_i \frac{\delta_n(\hat{f}(x_i) - f^*(x_i))}{||\hat{f} - f^*||_n} \\
&\leq 2\sigma \frac{||\hat{f} - f^*||_n}{\delta_n} Z_n(\delta_n)
\end{aligned}$$

where $Z_n(\delta_n) = \sup_{||f-f^*||_n \leq \delta_n} \left| \frac{1}{n} \sum_{i=1}^{n} w_i(f(x_i) - f^*(x_i)) \right|$ We have shown that

$$||\hat{f} - f^*||_n^2 \leq 4 \left\{ \frac{\sigma Z_n(\delta_n)}{\delta_n} \right\}^2$$

which holds for any kind of noise. Define

$$\epsilon(t) = \{w \mid Z_n(\delta_n) \leq \mathbb{E}[Z(\delta_n)] + t\delta_n\}$$

Thus, we get

$$\begin{aligned}
||\hat{f} - f^*||_n^2 &\leq 4 \left\{ \frac{\sigma(G_n(\delta_n) + t\delta_n)}{\delta_n} \right\}^2 \\
&= 4 \left\{ \sigma \frac{G_n(\delta_n)}{\delta_n} + \sigma t \right\}^2 \quad \text{with probability } 1 - \mathbb{P}(\epsilon^c(t)) \\
&= 4 \left\{ \frac{\delta_n}{2} + \sigma t \right\} \\
&= 4 \left\{ \frac{\delta_n}{2} + \sigma t \right\}^2 \\
&\leq 8 \left\{ \frac{\delta_n^2}{4} + \sigma^2 t^2 \right\}
\end{aligned}$$