3.6 Finite Hypothesis Classes

We now show the consequence of finite hypothesis classes plus concentration inequalities.

[Theorem 7] finite hypothesis class

Let H be a hypothesis class: $\mathcal{H} = \{h : \mathcal{X} \to \mathcal{Y}\}$.

Let ℓ be one-zero loss: $\ell((x,y),h) = \mathbb{1}\{y \neq h(x)\}$.

Assume $\mathcal{H}$ is finite, let $\hat{h}$ be the empirical risk minimizer.

Then with prob. $\geq 1-\delta$, the excess risk is bounded as follows:

$$L(\hat{h}) - L(h^*) \leq \sqrt{2(\log|H| + \log(2/\delta))/n} = O\left(\sqrt{\frac{\log|H|}{n}}\right)$$

Pf: Step 1: convergence

For a fixed $h \in \mathcal{H}$, note that $\hat{L}(h)$ is an empirical average over n i.i.d. loss terms (bounded in $[0,1]$) with expectation $L(h)$. Then by Hoeffding's inequality,

$$P\{\hat{L}(h) - L(h) \geq \varepsilon\} \leq \exp\{-2n\varepsilon^2\}$$

Apply the bound again on the negative loss then we have

$$P\{L(h) - \hat{L}(h) \geq \varepsilon\} \leq \exp\{-2n\varepsilon^2\}$$

then we have union bound:

$$P\{|\hat{L}(h) - L(h)| \geq \varepsilon\} \leq 2\exp\{-2n\varepsilon^2\}$$

Step 2: uniform convergence.

Apply union bound over $\mathcal{H}$ and we obtain:

$$P\{\sup_{h\in\mathcal{H}} |\hat{L}(h) - L(h)| \geq \tfrac{\varepsilon}{2}\} \leq |H| \cdot 2\exp\{-2n(\tfrac{\varepsilon}{2})^2\} := \delta$$

Therefore
$$\begin{aligned} L(\hat{h}) - L(h^*) &= L(h) - \hat{L}(h) + \hat{L}(h) - \hat{L}(h^*) + \hat{L}(h^*) - L(h^*) \\ &\leq L(h) - \hat{L}(h) + \hat{L}(h^*) - L(h^*) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{with prob.} \geq 1-\delta \quad \square \end{aligned}$$

Note: ① still depend on $|H|$ and $1/\delta$

② We get a $\frac{1}{\sqrt{n}}$ rate, while the realizable bound had a $\frac{1}{n}$.

③ When we perform asymptotics, we also got a $\frac{1}{n}$ rate. but this risk is w.r.t. the small area around θ^* whereas here we globally across $\mathcal{H}$.