Rademacher Complexity

$\sigma_i \pm 1$

$A \subset \mathbb{R}^m$

$$R(A) := \frac{1}{m} \mathop{\mathbb{E}}_{\sigma}\left[\sup_{a \in A} \sum \sigma_i a_i\right]$$

$$A = \ell \circ \mathcal{H} \circ S = \{(\ell(h, z_1), \dots, \ell(h, z_m)) : h \in \mathcal{H}\}$$

Rademacher Calculus

Lemma 1

For $A \subset \mathbb{R}^m$, $c \in \mathbb{R}$ and $a_0 \in \mathbb{R}^m$

$$R(\{ca + a_0 : a \in A\}) \leq |c| R(A)$$

Pf

$$\frac{1}{m} \mathop{\mathbb{E}}_{\sigma}\left[\sup_{a \in A} \sum \sigma_i (c a_i + a_{0i})\right] = \frac{1}{m} \mathop{\mathbb{E}}_{\sigma}\left[\cancelto{0}{\sum \sigma_i a_{0i}} + \sup_{a \in A} \sum \sigma_i a_i c\right]$$

$$= \frac{1}{m} \mathop{\mathbb{E}}_{\sigma}\left[\sup_{a \in A} \sum \sigma_i a_i c\right]$$

$$\leq |c| R(A)$$

Lemma 2

$$A \subset \mathbb{R}^m, \quad A' = \left\{\sum_{j=1}^{N} \alpha_j a^{(j)} : N \in \mathbb{N}, \forall j, a^{(j)} \in A, \alpha_j \geq 0, \|\alpha\|_1 = 1\right\}$$

Then,

$$R(A') = R(A)$$

Fact, $\forall v$

$$\sup_{\alpha \geq 0 : \|\alpha\|_1 = 1} \sum \alpha_j v_j = \max_j v_j$$

$$m R(A') = \mathop{\mathbb{E}}_{\sigma}\left[\sup_{a' \in A'} \sum \sigma_i a'\right] = \mathop{\mathbb{E}}_{\sigma}\left[\sup_{\substack{\alpha \geq 0 \\ \|\alpha\|_1 = 1 \\ a^1, \dots a^N \in A \\ N \in \mathbb{N}}} \sum_{i=1} \sigma_i \sum_{j=1}^{N} \alpha_j a_i^j\right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{\alpha \geq 0, \|\alpha\|_1 = 1} \quad \sup_{a^1, \ldots, a^N} \sum_{i=1}^{m} \sigma_i \sum_{j=1}^{N} \alpha_j a_i^{(j)} \right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{\alpha \geq 0, \|\alpha\|_1 = 1} \sum_{j=1}^{N} \alpha_j \sup_{a^j} \sum_{i=} \sigma_i a_i^j \right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{a \in A} \sum \sigma_i a_i \right] = m R(A)$$

Linear Classes

$\mathcal{H}_2 = \{ x \mapsto \langle w, x \rangle : \| w_2 \| \leq 1 \}$

Thm: $S = (x_1, \ldots, x_m)$ ~~Def~~^{n} $\mathcal{H}_2 \circ S = \{ (\langle w, x_1 \rangle, \ldots, \langle w, x_m \rangle) : \|w\|_2 \leq 1 \}$

$$R(\mathcal{H}_2 \circ S) \leq \frac{\max_i \|x_i\|_2}{\sqrt{m}}$$

Pf

$$m R(\mathcal{H}_2 \circ S) = \mathop{E}_{\sigma}\left[\sup_{a \in \mathcal{H}_2 \circ S} \sum^{m} \sigma_i a_i \right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{w : \|w\| \leq 1} \sum^{m} \sigma_i \langle w, x_i \rangle \right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{w : \|w\| \leq 1} \left\langle w, \sum^{m} \sigma_i x_i \right\rangle \right]$$

$$\leq \mathop{E}_{\sigma}\left[\left\| \sum^{m} \sigma_i x_i \right\|_2 \right]$$

$$= \mathop{E}_{\sigma}\left[\left(\left\|\sum^{m}\sigma_i x_i\right\|_2^2\right)^{\frac{1}{2}}\right] \le \left(\mathop{E}_{\sigma}\left[\left\|\sum \sigma_i x_i\right\|_2^2\right]\right)^{\frac{1}{2}}$$

$$= \left(\mathop{E}_{\sigma}\left[\sum_{i,j} \sigma_i\sigma_j\langle x_i, x_j\rangle\right]\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i\neq j}\langle x_i, x_j\rangle \mathop{E}_{\sigma}\left[\sigma_i\sigma_j\right]^{\nearrow 0} + \sum \langle x_i, x_i\rangle \mathop{E}_{\sigma}\left[\sigma_i^2\right]^{\nearrow 1}\right)^{1/2}$$

$$= \left(\sum \|x_i\|_2^2\right)^{1/2} \le \sqrt{m}\max_i \|x_i\|_2$$

~~Lemma 2 : (Massart) $a_i \in \mathbb{R}^m$~~

~~$A = \{a_1, \dots, a_N\}$ $N \in \mathbb{N}$~~

~~$\bar{a} := \frac{1}{N}\sum a_i$~~

Lemma 3 : (Massart)

$a_i \in \mathbb{R}^m$ $\quad \bar{a} := \frac{1}{N}\sum a_i \quad$ N finite

$A = \{a_1, \dots, a_N\}$

$$R(A) \le \max_{a\in A}\|a-\bar{a}\|\frac{\sqrt{2\log N}}{m}$$

Pf: from Lemma 1 wlog $\bar{a} = 0$

Let $\lambda > 0$ $A' = \{\lambda a_1, \dots, \lambda a_n\}$

$$m R(A') = \mathop{E}_{\sigma}\left[\max_{A'} \langle\sigma, a\rangle\right] = \mathop{E}_{\sigma}\left[\log\left(\max_{A'} e^{\langle\sigma, a\rangle}\right)\right]$$

$$\leq \mathop{E}_{\sigma}\left[\log\left(\sum_{A'} e^{\langle\sigma,a\rangle}\right)\right]$$

$$\leq \log\left(\mathop{E}_{\sigma}\left[\sum_{A'} e^{\langle\sigma,a\rangle}\right]\right)$$

$$= \log\left(\sum_{A'} \prod_{\sigma} \underbrace{\mathop{E}_{\sigma_i}\left[e^{\sigma_i a_i}\right]}_{}\right)$$

$$\frac{1}{2}e^{(a_i)} + \frac{1}{2}e^{-a_i} \leq e^{a_i^2/2}$$

$$\leq \log\left(\sum_{A'} \prod_{\sigma} e^{a_i^2/2}\right) = \log\left(\sum_{A'} e^{\|a_i\|^2/2}\right)$$

$$\leq \log\left(|A'| \max_{A'} e^{\|a\|^2/2}\right) = \log|A'| + \max_{A'} \|a\|^2/2$$

lemma 1 $\rightarrow$ $R(A) = \frac{1}{\lambda} R(A')$

$$R(A) = \frac{\log|A| + \lambda^2 \max_{A} \|a\|^2/2}{\lambda m}$$

Set $\lambda = \sqrt{2\log|A| / \max_{A} \|a\|^2}$

$$R(A) \leq \max_{A} \|a - \bar{a}\| \frac{\sqrt{2\log(N)}}{m}$$

Lemma 4: For each $i \in [m]$, $\phi_i : \mathbb{R} \to \mathbb{R}$ ℓ-lipschitz

$\Leftrightarrow$

$\alpha, \beta \in \mathbb{R}$ $|\phi_i(\alpha) - \phi_i(\beta)| \leq \ell |\alpha - \beta|$

For $a \in \mathbb{R}^m$, let $\phi(a)$ denote $(\phi_1(a_1), \dots, \phi_m(a_m))$

Let $\phi \circ A = \{\phi(a) : a \in A\}$

$$R(\phi \circ A) \leq \ell R(A)$$

Pf: let $\ell = 1$ (lemma 1) $\phi = \frac{1}{\ell}\phi'$

Let $A_i = \{(a_1, \dots, a_{i-1}, \phi_i(a_i), a_{i+1}, \dots, a_m) : a \in A\}$

Wlog $i = 1$

$$mR(A_1) = \mathop{E}_{\sigma}\left[\sup_{A_1} \sum^m \sigma_i a_i\right]$$

$$= \mathop{E}_{\sigma}\left[\sup_{A} \sigma_1 \phi_1(a_1) + \sum_2^m \sigma_i a_i\right]$$

$$= \frac{1}{2} \mathop{E}_{\sigma_2, \dots, \sigma_m}\left[\sup_A\left(\phi_1(a_1) + \sum_2^m \sigma_i a_i\right) + \sup_A\left(-\phi_1(a_1) + \sum_2^m \sigma_i a_i\right)\right]$$

$$= \frac{1}{2} \mathop{E}_{\sigma_2, \dots, \sigma_m}\left[\sup_{a, a'}\left(\phi_1(a_1) - \phi_1(a_1') + \sum_2^m \sigma_i a_i + \sum_2^m \sigma_i a_i'\right)\right]$$

$$\leq \frac{1}{2} \mathop{E}_{\sigma_2, \dots, \sigma_m}\left[\sup_{a, a'}\left(|a_1 - a_1'| + \sum_2^m \sigma_i a_i + \sum_2^m \sigma_i a_i'\right)\right]$$

$$= \frac{1}{2} \mathop{E}_{\sigma_2, \dots, \sigma_m}\left[\sup_{a, a'}\left(a_1 - a_1' + \sum_2^m \sigma_i a_i + \sum_2^m \sigma_i a_i'\right)\right]$$