Department of Computer Science
UNIVERSITY OF COLORADO **BOULDER**

Classification: Rademacher Complexity

Machine Learning: Jordan Boyd-Graber
University of Colorado Boulder
LECTURE 6

Slides adapted from Rob Schapire

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Content Questions

Administrivia Questions

Administrivia Questions

Administrivia Questions

Administrivia Questions

Administrivia Questions

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$ \tag{1} $$

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$\mathscr{R}_m(H) = \mathbb{E}_{S \sim D^m, \sigma}\left[\sup_{h \in H} \frac{1}{m} \sum_{i=1}^{m} \sigma_i h(z_i)\right] \tag{1}$$

$$\tag{2}$$

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$\mathscr{R}_m(H) = \mathbb{E}_{S \sim D^m, \sigma} \left[\sup_{h \in H} \frac{1}{m} \sum_{i=1}^{m} \sigma_i h(z_i) \right] \tag{1}$$

$$= \mathbb{E}_{S \sim D^m, \sigma} \left[\frac{1}{m} \sum_{i=1}^{m} \sigma_i h_0(z_i) \right] \tag{2}$$

$$\tag{3}$$

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$\mathscr{R}_m(H) = \mathbb{E}_{S \sim D^m, \sigma}\left[\sup_{h \in H} \frac{1}{m} \sum_{i=1}^{m} \sigma_i h(z_i)\right] \tag{1}$$

$$= \mathbb{E}_{S \sim D^m, \sigma}\left[\frac{1}{m} \sum_{i=1}^{m} \sigma_i h_0(z_i)\right] \tag{2}$$

$$= \mathbb{E}_{S \sim D^m}\left[\frac{1}{m} \sum_{i=1}^{m} \mathbb{E}_{\sigma}\left[\sigma_i\right] \sigma_i h_0(z_i)\right] \tag{3}$$

$$\tag{4}$$

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$\mathscr{R}_m(H) = \mathbb{E}_{S \sim D^m, \sigma}\left[\sup_{h \in H} \frac{1}{m}\sum_{i=1}^{m} \sigma_i h(z_i)\right] \tag{1}$$

$$= \mathbb{E}_{S \sim D^m, \sigma}\left[\frac{1}{m}\sum_{i=1}^{m} \sigma_i h_0(z_i)\right] \tag{2}$$

$$= \mathbb{E}_{S \sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m} \mathbb{E}_{\sigma}[\sigma_i] \sigma_i h_0(z_i)\right] \tag{3}$$

$$= \mathbb{E}_{S \sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m} 0 \cdot \sigma_i h_0(z_i)\right] \tag{4}$$

Single Hypothesis

What is the Rademacher complexity of a hypothesis set reduced to a single hypothesis?

$$\mathscr{R}_m(H) = \mathbb{E}_{S\sim D^m,\sigma}\left[\sup_{h\in H}\frac{1}{m}\sum_{i=1}^{m}\sigma_i h(z_i)\right] \tag{1}$$

$$= \mathbb{E}_{S\sim D^m,\sigma}\left[\frac{1}{m}\sum_{i=1}^{m}\sigma_i h_0(z_i)\right] \tag{2}$$

$$= \mathbb{E}_{S\sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m}\mathbb{E}_{\sigma}\left[\sigma_i\right]\sigma_i h_0(z_i)\right] \tag{3}$$

$$= \mathbb{E}_{S\sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m}0\cdot\sigma_i h_0(z_i)\right] = 0 \tag{4}$$

$$\tag{5}$$

Rademacher Identity 1

Prove

$$\mathscr{R}_m(\alpha H) = |\alpha| \mathscr{R}_m(H)$$

If $\alpha \geq 0$

If $\alpha < 0$

Rademacher Identity 1

Prove

$$\mathscr{R}_m(\alpha H) = |\alpha| \mathscr{R}_m(H)$$

If $\alpha \geq 0$

$$\sup_{h \in \alpha H} \sum_{i=1}^{m} \sigma_i h(x_i) = \tag{6}$$

$$\sup_{h \in H} \sum_{i=1}^{m} \alpha \sigma_i h(x_i) = \tag{7}$$

$$\alpha \sup_{h \in H} \sum_{i=1}^{m} \sigma_i h(x_i) \tag{8}$$

If $\alpha < 0$

Rademacher Identity 1

Prove

$$\mathscr{R}_m(\alpha H) = |\alpha| \mathscr{R}_m(H)$$

If $\alpha \geq 0$

$$\sup_{h \in \alpha H} \sum_{i=1}^{m} \sigma_i h(x_i) = \tag{6}$$

$$\sup_{h \in H} \sum_{i=1}^{m} \alpha \sigma_i h(x_i) = \tag{7}$$

$$\alpha \sup_{h \in H} \sum_{i=1}^{m} \sigma_i h(x_i) \tag{8}$$

If $\alpha < 0$

$$\sup_{h \in \alpha H} \sum_{i=1}^{m} \sigma_i h(x_i) = \tag{9}$$

$$\sup_{h \in H} \sum_{i=1}^{m} \alpha \sigma_i h(x_i) = \tag{10}$$

$$(-\alpha) \sup_{h \in H} \sum_{i=1}^{m} (-\sigma_i) h(x_i) \tag{11}$$

Rademacher Identity 1

Prove

$$\mathscr{R}_m(\alpha H) = |\alpha| \mathscr{R}_m(H)$$

If $\alpha \geq 0$

$$\sup_{h \in \alpha H} \sum_{i=1}^{m} \sigma_i h(x_i) = \tag{6}$$

$$\sup_{h \in H} \sum_{i=1}^{m} \alpha \sigma_i h(x_i) = \tag{7}$$

$$\alpha \sup_{h \in H} \sum_{i=1}^{m} \sigma_i h(x_i) \tag{8}$$

If $\alpha < 0$

$$\sup_{h \in \alpha H} \sum_{i=1}^{m} \sigma_i h(x_i) = \tag{9}$$

$$\sup_{h \in H} \sum_{i=1}^{m} \alpha \sigma_i h(x_i) = \tag{10}$$

$$(-\alpha) \sup_{h \in H} \sum_{i=1}^{m} (-\sigma_i) h(x_i) \tag{11}$$

Since σ_i and $-\sigma$ have the same distribution, $\mathscr{R}_m(\alpha H) = |\alpha| \mathscr{R}_m(H)$

Rademacher Identity 2

Prove

$$\mathscr{R}_m(H+H') = \mathscr{R}_m(H) + \mathscr{R}_m(H')$$

(12)

Rademacher Identity 2

Prove

$$\mathscr{R}_m(H + H') = \mathscr{R}_m(H) + \mathscr{R}_m(H')$$

$$\mathscr{R}_m(H + H') \tag{12}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h \in H, h' \in H'} \sum_{i=1}^{m} \sigma_i (h(x_i) + h'(x_i))\right] \tag{13}$$

$$\tag{14}$$

Rademacher Identity 2

Prove

$$\mathscr{R}_m(H+H') = \mathscr{R}_m(H) + \mathscr{R}_m(H')$$

$$\mathscr{R}_m(H+H') \tag{12}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i (h(x_i) + h'(x_i))\right] \tag{13}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i h(x_i) + \sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i h'(x_i))\right] \tag{14}$$

(15)

Rademacher Identity 2

Prove

$$\mathscr{R}_m(H+H') = \mathscr{R}_m(H) + \mathscr{R}_m(H')$$

$$\mathscr{R}_m(H+H') \tag{12}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i (h(x_i) + h'(x_i))\right] \tag{13}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i h(x_i) + \sup_{h\in H, h'\in H'} \sum_{i=1}^{m} \sigma_i h'(x_i))\right] \tag{14}$$

$$= \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h\in H} \sum_{i=1}^{m} \sigma_i h(x_i)\right] + \frac{1}{m}\mathbb{E}_{\vec{\sigma},S}\left[\sup_{h'\in H'} \sum_{i=1}^{m} \sigma_i h(x_i)\right] \tag{15}$$