# Rademacher复杂度

计算机学院 20020129 王悟信

定理：对假设空间$\mathcal{H}:\chi\rightarrow\{-1,+1\}$，根据分布$\mathcal{D}$从$\chi$中独立同分布采样得到示例集$\mathcal{D}=x_1,x_2,\cdots,x_m,x_i\in\chi,0<\delta<1$，对任意$h\in\mathcal{H}$，以至少$1-\delta$的概率有

$$E(h)\leq\hat{E}(h)+R_m(\mathcal{H})+\sqrt{\frac{ln(1/\delta)}{2m}},\tag{12.47}$$

$$E(h)\leq\hat{E}(h)+\hat{R}_D(\mathcal{H})+3\sqrt{\frac{ln(2/\delta)}{2m}},\tag{12.48}$$

证明：

> - 引理
>   实值函数空间$\mathcal{F}:\mathcal{Z}\rightarrow[0,1]$，根据分布$\mathcal{D}$从$\mathcal{Z}$中独立同分布采样得到数据集$\mathcal{Z}=\{x_1,x_2,\cdots,x_m\}$，$x_i\in\mathcal{Z},0<\delta<1$，对任意$f\in\mathcal{F}$，以至少$1-\delta$的概率有
>
> $$\mathbb{E}[f(x)]\leq\frac{1}{m}\sum_{i=1}^{m}f(x_i)+2R_m(\mathcal{F})+\sqrt{\frac{ln(1/\delta)}{2m}}$$
>
> $$\mathbb{E}[f(x)]\leq\frac{1}{m}\sum_{i=1}^{m}f(x_i)+2\hat{R}_Z(\mathcal{F})+3\sqrt{\frac{ln(2/\delta)}{2m}}$$

下面开始证明：

对二分类问题的假设空间$\mathcal{H}$，令$\mathcal{Z}=\chi\times\{-1,+1\}$，则$\mathcal{H}$中的假设变形为

$$f_h(z)=f_h(x,y)=\mathbb{I}(h(x)\neq y)$$

于是就可以将值域为$\{-1,+1\}$的假设空间$\mathcal{H}$转化为值域为$[0,1]$的函数空间$\mathcal{F}=\{f_h:h\in\mathcal{H}\}$

则

$$
\begin{aligned}
\hat{R}_Z(\mathcal{F}_{\mathcal{H}}) &= \mathbb{E}_{\sigma}[\sup_{f_h \in \mathcal{F}_{\mathcal{H}}} \frac{1}{m}\sum_{i=1}^{m}\sigma_i f_h(x_i, y_i)] \\
&= \mathbb{E}_{\sigma}[\sup_{h \in \mathcal{H}} \frac{1}{m}\sum_{i=1}^{m}\sigma_i \mathbb{I}(h(x_i) \neq y_i)] \\
&= \mathbb{E}_{\sigma}[\sup_{h \in \mathcal{H}} \sum_{i=1}^{m}\sigma_i \frac{1 - y_i h(x_i)}{2}] \\
&= \frac{1}{2}\mathbb{E}_{\sigma}[\frac{1}{m}\sum_{i=1}^{m}\sigma_i + \sup_{h \in \mathcal{H}} \frac{1}{m}\sum_{i=1}^{m}(-y_i \sigma_i h(x_i))] \\
&= \frac{1}{2}\mathbb{E}_{\sigma}[\sup_{h \in \mathcal{H}} \frac{1}{m}\sum_{i=1}^{m}(-y_i \sigma_i h(x_i))] \\
&= \frac{1}{2}\mathbb{E}_{\sigma}[\sup_{h \in \mathcal{H}} \frac{1}{m}\sum_{i=1}^{m}(\sigma_i h(x_i))] \\
&= \frac{1}{2}\hat{R}_D(\mathcal{H})
\end{aligned}
$$

对上式求期望可得：

$$
\mathbb{E}_Z[\hat{R}_Z(\mathcal{F}_{\mathcal{H}})] = \frac{1}{2}\mathbb{E}_Z[\hat{R}_D(\mathcal{H})]
$$

而

$$
\mathbb{E}_Z[\hat{R}_Z(\mathcal{F}_{\mathcal{H}})] = R_m(\mathcal{F}_{\mathcal{H}})
$$

$$
\mathbb{E}_Z[\hat{R}_D(\mathcal{H})] = R_m(\mathcal{H})
$$

故

$$
R_m(\mathcal{F}_{\mathcal{H}}) = \frac{1}{2}R_m(\mathcal{H})
$$

根据引理，将$h$即为$f(x)$，带入可得

对于任意$h \in \mathcal{H}$，以至少$1 - \delta$的概率有

$$
\mathbb{E}(h) \leq \hat{E}(h) + 2R_m(\mathcal{F}_{\mathcal{H}}) + \sqrt{\frac{ln(1/\delta)}{2m}}
$$

$$\mathbb{E}(h) \leq \hat{E}(h) + 2\hat{R}_Z(\mathcal{F}_\mathcal{H}) + 3\sqrt{\frac{ln(2/\delta)}{2m}}$$

而，根据推理可得

将$R_m(\mathcal{F}_\mathcal{H})$替换为$\frac{1}{2}R_m(\mathcal{H})$

将$\hat{R}_Z(\mathcal{F}_\mathcal{H})$替换为$\frac{1}{2}\hat{R}_D(\mathcal{H})$

则该定理得证。