Machine Learning Homework: Week 3 & 4

周书予

2000013060@stu.pku.edu.cn

March 20, 2022

1

Statement

Prove that

$$\begin{aligned}\frac{1}{2}\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\mathbb{E}_z\phi(z)-\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)\right|\geqslant 2\varepsilon\right)&\leqslant\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)\right|\geqslant\varepsilon\right)\\&\leqslant 2\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\mathbb{E}_z\phi(z)-\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)\right|\geqslant\frac{\varepsilon}{2}\right)\end{aligned}\tag{1}$$

Solution

$$\begin{aligned}&\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)\right|\geqslant\varepsilon\right)\\=&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)\right|>\varepsilon-\delta\right)\\\leqslant&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>\frac{\varepsilon-\delta}{2}\text{ or }\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>\frac{\varepsilon-\delta}{2}\right)\\\leqslant&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>\frac{\varepsilon-\delta}{2}\right)+\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>\frac{\varepsilon-\delta}{2}\right)\\=&\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|\geqslant\frac{\varepsilon}{2}\right)+\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|\geqslant\frac{\varepsilon}{2}\right)\\=&2\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\mathbb{E}_z\phi(z)-\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)\right|\geqslant\frac{\varepsilon}{2}\right)\end{aligned}\tag{2}$$

which provides the proof for the inequality on the right side.

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)\right|\geqslant\varepsilon\right)\\
=&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)\right|>\varepsilon-\delta\right)\\
\geqslant&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>2(\varepsilon-\delta)\text{ and }\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|<\varepsilon-\delta\right)\\
\geqslant&\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>2(\varepsilon-\delta)\right)\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|<\varepsilon-\delta\right)
\end{aligned}
\tag{3}
$$

By Chernoff bound we obtain that $\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=n+1}^{2n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|<\varepsilon\right)\geqslant\frac{1}{2}$ when $n\geqslant\frac{\ln 2}{\varepsilon^2}$, which gives the proof for the inequality on the left side

$$
\begin{aligned}
\cdots\geqslant&\frac{1}{2}\mathbb{P}\left(\forall\delta>0,\exists\phi\in\Phi\text{ s.t. }\left|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)-\mathbb{E}_z\phi(z)\right|>2(\varepsilon-\delta)\right)\\
=&\frac{1}{2}\mathbb{P}\left(\sup_{\phi\in\Phi}\left|\mathbb{E}_z\phi(z)-\frac{1}{n}\sum_{i=1}^{n}\phi(z_i)\right|\geqslant 2\varepsilon\right)
\end{aligned}
\tag{4}
$$

2

Statement

Prove that

$$
\sum_{k=0}^{d}\binom{n}{k}<\left(\frac{\mathrm{e}n}{d}\right)^d
$$

for any $n\in\mathbb{N}^+$ and $1\leqslant d\leqslant n$.

Solution

By Taylor expansion we have

$$
\mathrm{e}^d=\sum_{k=0}^{\infty}\frac{d^k}{k!}>\sum_{k=0}^{d}\frac{d^k}{k!}
$$

where immediately follows the proof

$$
\begin{aligned}
\left(\frac{\mathrm{e}n}{d}\right)^d&>n^d\sum_{k=0}^{d}\frac{d^{k-d}}{k!}\\
&\geqslant n^d\sum_{k=0}^{d}\frac{n^{k-d}}{k!}\\
&=\sum_{k=0}^{d}\frac{n^k}{k!}\\
&\geqslant\sum_{k=0}^{d}\frac{n^{\underline{k}}}{k!}\\
&=\sum_{k=0}^{d}\binom{n}{k}
\end{aligned}
\tag{5}
$$

3

Statement

Prove that the following two programming problems are equivalent. For simplicity, we assume that the first problem has non-trivial solution, i.e. there exists $t > 0$ satisfying $y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b) \geqslant t$ $(\forall i \in [n])$.

$$\begin{aligned} \max_{\mathbf{w},b,t} \quad & t \\ \text{s.t.} \quad & y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b) \geqslant t \ (\forall i \in [n]) \\ & \|\mathbf{w}\| = 1 \end{aligned} \qquad \begin{aligned} \min_{\mathbf{w},b,t} \quad & \frac{1}{2}\|\mathbf{w}\| \\ \text{s.t.} \quad & y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b) \geqslant 1 \ (\forall i \in [n]) \end{aligned}$$

Solution

Let the solution to the two problems above be t_0 and $\mathbf{w}_0$ respectively. By the assumption, we have $t_0 > 0$.

By dividing the inequality in the first problem by t_0 we obtain $y_i\left(\frac{\mathbf{w}^{\mathrm{T}}}{t_0}\mathbf{x}_i + \frac{b}{t_0}\right) \geqslant 1$, so $\mathbf{w}' = \frac{\mathbf{w}}{t_0}$ is a candidate for the second problem, and thus $\|\mathbf{w}_0\| \leqslant \|\mathbf{w}'\| = \frac{1}{t_0}$.

By dividing the inequality in the second problem by $\|w_0\|$ we obtain $y_i\left(\frac{\mathbf{w}_0}{\|\mathbf{w}_0\|}\mathbf{x}_i + \frac{b}{\|\mathbf{w}_0\|}\right) \geqslant \frac{1}{\|\mathbf{w}_0\|}$, thus $t \geqslant \frac{1}{\|\mathbf{w}_0\|}$ since $\left\|\frac{\mathbf{w}_0}{\|\mathbf{w}_0\|}\right\| = 1$.

So we know that $\|\mathbf{w}_0\|t_0 = 1$, which shows the equivalence relationship between these two programming problems.