Machine Learning HW#1

1. From definition, Let $A = \{\alpha_1, \dots \alpha_d \in \mathbb{Z} : \alpha_i \geq 0, \sum_{i=1}^{d} \alpha_i = k\}$

Then $|A| = |\mathcal{M}_{d,k}|$. Let $B = \{\beta_1 \dots \beta_d \in \mathbb{Z} : \beta_i \geq 1, \sum_{i=1}^{d} \beta_i = k+d\}$

$s_1\, \text{O} \mid s_2\, \text{O} \; \dots \; \text{O} \mid \text{O}\, s_{k+d}$ We have $|A| = |B|$ (by natural bijection)

By standard combinatorics, $|B|$ is the way to put $d-1$ blocks in the $k+d-1$ crevices of $k+d$ balls, which by definition is $\binom{k+d-1}{d-1}$

Thus $|A| = |B| = \binom{k+d-1}{d-1}$

Therefore $|\mathcal{M}_{d,k}| = \binom{d+k-1}{d-1}$

$$\binom{d+k-1}{d-1} = \frac{(d+k-1)!}{(d-1)!\,k!} = \frac{(d+k-1)\cdots(k+1)}{(d-1)!} = \frac{(d-1+k)\cdots(1+k)}{(d-1)\cdots 1} = \prod_{j=1}^{d-1} \frac{k+j}{j}$$

$$\geq \left(\frac{k+d-1}{d-1}\right)^{d-1} \quad \left(\text{As when } 1 \leq j \leq d-1, \; \frac{k+j}{j} \geq \frac{k+d-1}{d-1}\right)$$

$$= \left(1 + \frac{k}{d-1}\right)^{d-1} \quad \#$$

(See next page)

2.

(a) Observe that $\{x \in V_S^c\} = \{x$ and x_1 is not in same cell$\} \cap \{x$ and X_2 not in same cell$\} \cap \cdots \cap \{x$ and x_n not in same cell$\}$

Noticing $x, x_1 \cdots x_d$ are independent, $Pr(x \in V_S^c) = \prod_{i=1}^{n} Pr(x$ and x_i not in same cell$)$

$= Pr(x$ and x_1 not in same cell$)^n \leq (1 - \frac{1}{\# \text{cells}})^n = (1-h^d)^n$

Thus $\mathbb{E}_{S,x}(Pr(x \in V_S^c)) = \mathbb{E}_{S,x}\ (1-h^d)^n = (1-h^d)^n$. #

(b) Lemma: $\forall t \in (0,1)$, $(1-t)^n \leq \frac{1}{nt}$

Proof. $nt(1-t)^n = (nt)(1-t)\cdots(1-t) \leq \left(\frac{nt+1-t+\cdots+1-t}{n+1}\right)^{n+1} = \left(\frac{n}{n+1}\right)^{n+1} < 1$

by AM-GM inequality #

We then have

$$\inf_{h \in (0,1)} [h + (1-h^d)^n] \leq \inf_{h \in (0,1)} [h + \frac{1}{nh^d}]$$

According to AM-GM inequality,

$h + \frac{1}{nh^d} = \underbrace{\frac{h}{d} + \cdots + \frac{h}{d}}_{d \text{ terms}} + \frac{1}{nh^d} \geq \sqrt[d+1]{\frac{1}{d^d}\frac{1}{n}}$, equality holds when $h = \sqrt[d+1]{\frac{d}{n}}$

Thus $\inf_{h \in (0,1)} (h + \frac{1}{nh^d}) = C_1 n^{-\frac{1}{d+1}}$, where $C_1 = \sqrt[d+1]{\frac{1}{d^d}}$

Therefore $\min_{h \in [0,1]} [h + (1-h^{-d})^n] \leq \inf_{h \in (0,1)} [h + (1-h^{-d})^n] \leq C_1 \frac{1}{n^{\frac{1}{d+1}}}$ #.

(See next page)

(c) Lemma: if $x \in V_s$, $\|x - T(x)\| \le \sqrt{d}\, h$

Fix $h = \frac{1}{H}$. Proof. $\|x - T(x)\|_2 = \sqrt{(x_1 - T(x)_1)^2 + \cdots (x_d - T(x)_d)^2} \le \sqrt{d\, h^2} = h\sqrt{d}$

according to definition of V_s. #

On the other hand, $\forall x$, $\|x - T(x)\|_2 \le \sqrt{1^2 + \cdots + 1^2} = \sqrt{d}$

Noticing $\|x - T(x)\| = \|x - T(x)\| \mathbb{1}_{x \in V_s} + \|x - T(x)\| \mathbb{1}_{x \in V_s^c}$

We have

$$\mathbb{E}_{S,x} \|x - T(x)\| = \mathbb{E}_{S,x} \|x - T(x)\| \mathbb{1}_{x \in V_s} + \mathbb{E}_{S,x} \|x - T(x)\| \mathbb{1}_{x \in V_s^c}$$

$\le \sqrt{d}\, h \cdot 1 + \sqrt{d}\,(1 - h^d)^n = \sqrt{d}\,(h + (1 - h^d)^n)$ for all $h = \frac{1}{H}$

Because $\mathbb{E}_{S,x} \|x - T(x)\|_2$ is independent of the choice of H,

We have $\mathbb{E}_{S,x} \|x - T(x)\|_2 \le \sqrt{d} \min_{h \in [0,1]} h + (1 - h^d)^n \le C_1 \sqrt{d} \frac{1}{n^{\frac{1}{d+1}}}$ #.

Given S,

(d) $\|\hat{f}_s - f^*\|_{L^1} = \int_{[0,1]^d} |\hat{f}_s(x) - f^*(x)|\, d\rho = \int_{[0,1]^d} |f^*(T(x)) - f^*(x)|\, d\rho$

$\le \int_{[0,1]^d} L \|T(x) - x\|_2\, d\rho = L\, \mathbb{E}_{x \sim \rho} \|T(x) - x\|_2$

First "=" is definition of L^1 norm, second "=" is def of $\hat{f}_s(x)$, "$\le$" is the given condition, Third "=" is def of $\mathbb{E}_{x \sim \rho}[\cdot]$.

Therefore $\mathbb{E}_S \|\hat{f}_s - f^*\|_{L^1} \le L\, \mathbb{E}_S\, \mathbb{E}_{x \sim \rho} \|T(x) - x\|_2$

$$= L\, \mathbb{E}_{S,x} \|T(x) - x\|_2$$

$$\le L C_1 \frac{\sqrt{d}}{n^{\frac{1}{d+1}}}$$

$$\mathbb{E}_S \|\hat{f}_s - f^*\|_{L^1} \lesssim \frac{\sqrt{d}}{n^{\frac{1}{d+1}}}$$ #.