Homework 5: Mathematics for AI
Due date: December 9th, 11:59 pm.
Note: Submit your solution file to TA, Seojin Kim (osikjs@kaist.ac.kr), by email.
For questions, contact TA as well.

---

1. The Fundamental Theorem of Markov chains says that for a connected Markov chain, the long-term average distribution $\mathbf{a(t)}= \frac{1}{t}(\mathbf{p(0)} + \mathbf{p(1)} + ... + \mathbf{p(t\text{-}1)})$ converges to a stationary distribution. Does the t step distribution $\mathbf{p(t)}$ also converges for every connected Markov Chain? Prove or disprove it.

2. Consider a 2-dimensional lattice with n vertices in each coordinate. The transition probability of each neighbor is $\frac{1}{4}$. Each boundary point has a self-loop with transition probability of $1 - \frac{\text{(number of neighbors)}}{4}$. Find the ε-mixing time for this graph in the following way:

* "Edges leaving S" means the edges s.t. one vertex is in S and the other vertex is not in S.

(a) Show that the minimum number of edges leaving a set S is n when $\frac{n^2}{2} \geq |S| \geq \frac{n^2}{4}$.

(b) Show that the minimum number of edges leaving a set S is $\lfloor 2\sqrt{|S|} \rfloor$ when $|S| \leq \frac{n^2}{4}$.

(c) Compute Φ.

(d) Compute ε-mixing time.

3. Let S be a set of examples, and $\mathcal{H}$ be a hypothesis class. To prove that the VC-dimension of $\mathcal{H}$ is some integer d, we must show: (1) there exists a subset of S of size d shattered by $\mathcal{H}$, and (2) there exist no subsets of S of size $D \geq d$ that are shattered by $\mathcal{H}$.

(a) Prove that if there exists a subset of S of size d shattered by $\mathcal{H}$, then for any $1 \leq k < d$ there also exists a subset of size k of X shattered by $\mathcal{H}$.

* By proving (a), we can relax (2) to the following statement: there exist no subsets S of size $d + 1$ that are shattered by $\mathcal{H}$.

(b) Let $S = \mathbb{R}$, and $\mathcal{H}$ be the set of all classifier h that classify a point x as $h(x) = 1$ if $x \in \cup_{i=1}^{p} R_i$, and $h(x) = -1$ otherwise for some set of non-intersecting intervals $R_1, ..., R_p$ with fixed given p. Find the VC-dimension of $\mathcal{H}$.

4. For $k = 1, 2$, let r_k be some positive integer and $\mathcal{F}_k : \{-1, 1\}^{r_k} \to \{-1, 1\}$ be some function class with VC-dimension d_k. Define $\mathcal{G}_k : \{-1, 1\}^{r_k} \to \{-1, 1\}^{r_{k+1}}$ with $r_3 = 1$ as:

$$\mathcal{G}_k = \{h(x) = (f_1(x), ..., f_{r_{k+1}}(x)) | f_1, ..., f_{r_{k+1}} \in \mathcal{F}_k\}.$$

Our hypothesis class $\mathcal{H}$ is 2-layer feedforward network defined by:

$$\mathcal{H} = \{h_2 \circ h_1 : \{-1, 1\}^{r_1} \to \{-1, 1\} | h_1 \in \mathcal{G}_1, h_2 \in \mathcal{G}_2\}.$$

(a) Prove that the growth function of $\mathcal{H}$ is bounded as:

$$\mathcal{H}[n] \leq F_1[n]^{r_2} \cdot F_2[n]$$

(b) Prove $\mathcal{H}[n] \leq (en)^d$, where $d = r_2 d_1 + d_2$.

Hint: Use Sauer's Lemma.

(c) Show that VC-dimension of $\mathcal{H}$ is $O(d \ln d)$.

Hint: Use the result of (b). When $n \geq 16,$ we have $\log_2 n \leq \sqrt{n}$.