Machine Learning Homework5

LiWei Wang

October 24, 2016

1 Problem

Figure out the relationship between Φ and $\mathcal{H}$.

2 Poblem

Compute the VC-Dimension of Linear Classifier.

3 Problem

Given a matrix $A = (a_{ij})_{n\times m}$, show that $\min_{i\leq n}\max_{j\leq m} a_{ij} \geq \max_{i\leq m}\min_{j\leq n} a_{ji}$.

4 Optional Problem, VC-inequality

Let p be a distribtuion over $[n]$ then let H be a family of subsets of $[n]$. Suppose the correpsonding family of indicator functions $F = \{I_S : S \in H\}$ has VC-dimension d. Independently take m samples from p, denoted by $X_1, X_2, ..., X_m$.

4.1

Prove that,

$$\mathbf{E}[\sup_{S\in H} |\frac{1}{m}\sum_{i=1}^{m} I[X_i \in S] - S(p)|] = O(\sqrt{\frac{d}{m}})$$

Where $S(p) = \sum_{i\in S} p_i$.

4.2

Show that if $m = O(\frac{n+\log\frac{1}{\delta}}{\epsilon^2})$ then with probability at least $1-\delta$, the L_1-distance between the empirical distribution $\frac{1}{m}\sum_{i=1}^{m}\delta_{X_i}$ and p is less than ϵ. Where δ_{X_i} is the Dirac delta function.

4.3

The Kolmogorov's distance between two distributions p and q is $\max_{i \leq n} |p(\{1, ..., i\}) - q(\{1, ..., i\})|$, i.e. the largest discrepency between their CDFs. Such that, if $m = O(\frac{\log \frac{1}{\delta}}{\epsilon^2})$ then with probability at least $1 - \delta$, the Kolmogorov's distance between $\frac{1}{m} \sum_{i=1}^{m} \delta_{X_i}$ and p is less than ϵ.