Partner 1
Partner 2
Partner 3

CSCI 1550 / 2540
May 3rd, 2024

Homework 4

Due: May 3rd, 2024 at 2:20pm
Remember to show your work for each problem to receive full credit.

Problem 1

Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be two range spaces. Prove that if $\mathcal{R}' \subseteq \mathcal{R}$ then the VC dimension of S' is no larger than the VC dimension of S.

Problem 2

Consider a random graph $G_{n,p}$ with n vertices and each pair of vertices independently connected by an edge with probability p. We prove a threshold for the existence of triangles in the graph.

Let $t_1, \ldots, t_{\binom{n}{3}}$ be an enumeration of all triplets of three vertices in the graph. Let $X_i = 1$ if the three edges of the triplet t_i appear in the graph, so that t_i forms a triangle in the graph. Otherwise $X_i = 0$. Let $X = \sum_{i=1}^{\binom{n}{3}} X_i$.

(a) Compute $E[X]$.

(b) Use (a) to show that if p is $o(\frac{1}{n})$, then $\lim_{n\to\infty} \Pr(X > 0) = 0$. (Note that we are using the small o-notation here.)

(c) Show that $Var[X_i] \leq p^3$.

(d) Show that $Cov(X_i, X_j) = p^5 - p^6$ for $O(n^4)$ pairs $i \neq j$, otherwise $Cov(X_i, X_j) = 0$.

(e) Show that $Var[X] = O(n^3p^3 + n^4p^5)$.

(f) Conclude that if p is $\omega(\frac{1}{n})$, then $\lim_{n\to\infty} \Pr(X = 0) = 0$.

Problem 3

We can generalize the problem of finding a large cut in a graph to finding a large k-cut. A k-cut is a partition of the vertices into k disjoint sets, and the value of a cut is the number of all edges crossing from one of the k sets to another.

(a) In class (also in section 6.2.1 in the book), we considered 2-cuts, and showed using the probabilistic method that any graph G with m edges has a cut with value at least $m/2$. Generalize this argument to show that any graph G with m edges has a k-cut with at value at least $(k-1)m/k$.

(b) Show how to use the conditional expectation argument (as we did in class, as described in section 6.3), to give a deterministic algorithm for finding such a cut.

Problem 4

We use the method of conditional probabilities to design a deterministic algorithm for the Set Balancing problem. Let A be an $n \times n$ matrix with entries in $\{0, 1\}$ and let $\bar{b}$ be an n-vector with entries in $\{-1, 1\}$. Let $\bar{c} = A\bar{b}$. Our goal is to minimize $||A\bar{b}||_\infty = \max_{1 \le i \le n} |c_i|$. We showed before that if $\bar{b}$ is chosen uniformly at random then

$$\Pr(\|A\bar{b}\|_\infty \ge \sqrt{4n \log n}) \le \frac{2}{n}.$$

Consider the following algorithm:

- Set $b_1 = 1$
- for $i = 2$ to n:
  - Let $\hat{b}_1, \ldots, \hat{b}_{i-1}$ be the assignments to $b_1, \ldots, b_{i-1}$.
  - Choose $\hat{b}_i \in \{-1, 1\}$ that minimizes $\max_{1 \le k \le n} \left| \sum_{j=1}^{i} a_{k,j} \hat{b}_j \right|$.

(a) Prove that if

$$\Pr(\|A\bar{b}\|_\infty \ge \sqrt{4n \log n} \mid \hat{b}_1 \ldots, \hat{b}_{i-1}) \le \frac{2}{n},$$

then there must be an assignment $\hat{b}_i$, such that

$$\Pr(\|A\bar{b}\|_\infty \ge \sqrt{4n \log n} \mid \hat{b}_1 \ldots, \hat{b}_i) \le \frac{2}{n}.$$

(b) Use conditional probabilities to prove that the above algorithm gives the required result, i.e., if $\bar{b} = [\hat{b}_1, \ldots, \hat{b}_n]^T$, then

$$\|A\bar{b}\|_\infty \le \sqrt{4n \log n}.$$

Homework 4

Problem 5

Suppose we are given an n-vertex undirected graph $G = (V, E)$, where the vertices are numbered $1, 2, \ldots, n$. An independent set of G is a subset $W \subseteq V$ of the vertices such that there is not an edge $e \in E$ connecting two vertices in W. We wish to find a (hopefully large) independent set in G. Given a permutation σ of the vertices, define a subset $S(\sigma)$ of the vertices as follows: for each vertex i, $i \in S(\sigma)$ if and only if no neighbor j of i precedes i in the permutation σ.

(a) Show that each $S(\sigma)$ is an independent set in G.

(b) Suggest a natural randomized algorithm to produce σ, for which you can show that the expected cardinality of $S(\sigma)$ is

$$\sum_{i=1}^{n} \frac{1}{d_i + 1}$$

where d_i denotes the degree of vertex i.

(c) Prove that G has an independent set of size at least $\sum_{i=1}^{n} \frac{1}{d_i+1}$.