MCS 549 – Foundations of Data Science
Fall 2026
Problem Set 2

Lev Reyzin

Problems

Prove all your answers.

1. Join two disjoint n-vertex cliques ($n \geq 3$) by one edge. Consider the lazy random walk: stay put with probability $1/2$; otherwise choose a neighbor uniformly. Find its stationary distribution π and prove that its conductance satisfies $\Phi = \Theta(n^{-2})$.

2. For a graph on $n \geq 1$ vertices and $\lambda > 0$, let $\pi(I) = \lambda^{|I|}/Z$ on independent sets (sets with no adjacent vertices), where $Z = \sum_{J \text{ independent}} \lambda^{|J|}$. A chain stays put with probability $1/2$; otherwise it chooses a vertex uniformly and proposes toggling its membership in I. Invalid proposals are rejected. Specify Metropolis–Hastings acceptance probabilities without using Z, and justify your answer.

3. Find the exact expected time for a simple random walk to visit every vertex of (a) a star with $n \geq 1$ leaves, starting at its center, and (b) the complete graph K_{n+1}, starting at any vertex. Express both answers using $H_n = \sum_{i=1}^{n} 1/i$, and compare them asymptotically. *Hint:* Condition on the number of vertices already visited.

4. Let $X = \{a, b\}$ and $\mathcal{H} = \{h_0, h_1\}$, where $h_0(a) = h_0(b) = h_1(a) = 0$ and $h_1(b) = 1$. A possibly randomized learner receives m independent labeled examples and must learn either target under every distribution on X. Fix $0 < \varepsilon, \delta < 1/4$.

i. For $D(b) = 2\varepsilon$ and $D(a) = 1 - 2\varepsilon$, show that for at least one target the probability of returning a hypothesis with error greater than ε is at least $\frac{1}{2}(1 - 2\varepsilon)^m$. *Hint:* What can the learner distinguish if b never appears?

ii. Deduce that guaranteeing error at most ε with probability at least $1 - \delta$ for both targets requires $m = \Omega(\log(1/\delta)/\varepsilon)$. Give a learner achieving a matching upper bound for every distribution on X.

5. For an integer $k \geq 1$, let $\mathcal{H}_k$ consist of unions of at most k intervals in $\mathbb{R}$. Prove that its VC dimension is $2k$, give an efficient consistent learner for realizable samples, and bound its PAC sample complexity in terms of k, ε, δ. You may cite the realizable VC generalization theorem.

6. Label the four points $x \in \{-1, 1\}^2$ by $y = x_1 x_2$. Prove that they are not affinely separable in $\mathbb{R}^2$. Find a feature map $\phi : \mathbb{R}^2 \to \mathbb{R}^3$ satisfying $K(x, z) = (x \cdot z)^2 = \phi(x) \cdot \phi(z)$, and exhibit a hyperplane separating the mapped points.