CS264: Homework #5

Due by midnight on Thursday, February 16, 2017

Instructions:

(1) Form a group of 1-3 students. You should turn in only one write-up for your entire group. See the course site for submission instructions.

(2) Please type your solutions if possible and feel free to use the LaTeX template provided on the course home page.

(3) **Students taking the course pass-fail should do five exercises and do not need to do any problems. Students taking the course for a letter grade should do six exercises, and we'll grade the Problems out of a total of 40 points.** Any points you receive on the problems in excess of 40 will be treated as extra-credit points.

(4) Write convincingly but not excessively. Exercise solutions rarely need to be more than 1-2 paragraphs. Problem solutions rarely need to be more than a half-page (per part), and can often be shorter.

(5) You may refer to your course notes, and to the textbooks and research papers listed on the course Web page *only*. You cannot refer to textbooks, handouts, or research papers that are not listed on the course home page. (Exception: feel free to use your undergraduate algorithms textbook.) Cite any sources that you use, and make sure that all your words are your own.

(6) If you discuss solution approaches with anyone outside of your team, you must list their names on the front page of your write-up.

(7) Exercises are worth 5 points each. Problem parts are labeled with point values.

(8) No late assignments will be accepted.

Lecture 9 Exercises

Exercise 24

Consider a random graph G drawn from the Erdös-Renyi distribution $\mathcal{G}(n, \frac{1}{2})$ described in lecture. Prove that, for every $\epsilon > 0$, the number of edges crossing every bisection of G is between $(1-\epsilon)\frac{n^2}{8}$ and $(1+\epsilon)\frac{n^2}{8}$ with probability $1 - o(1)$, where the $o(1)$ term tends to 0 as $n \to \infty$.

[Hint: this follows from the classic one-two punch of the Chernoff bound and the Union bound.]

Exercise 25

Let M be a symmetric $n \times n$ matrix. By the spectral theorem, M has an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of eigenvectors, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Prove the variational characterization of eigenvalues: for every $i \in \{1, 2, \ldots, n\}$,

$$\lambda_i = \max_{\mathbf{u} \in U_i} \mathbf{u}^T M \mathbf{u},$$

where U_i denotes the unit vectors that are orthogonal to $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}$.

Exercise 26

Let $G = (V, E)$ be an undirected graph and M the corresponding $V \times V$ adjacency matrix.

(a) Prove that the first eigenvalue λ_1 of M is at least the average degree of G.

(b) Prove that if G is d-regular (i.e., every vertex has exactly d neighbors), then $\lambda_1 = d$ and $|\lambda_i| \leq d$ for every $i \in \{1, 2, \ldots, n\}$.

[Hint: note that the operator $\mathbf{M}$ (sending $\mathbf{v} \mapsto \mathbf{M}\mathbf{v}$) has the following effect: for every vertex $i \in V$, the vertex's old value v_i is replaced by the sum of its neighbors' old values.]

Exercise 27

Continuing the previous exercise, let $G = (V, E)$ be a d-regular undirected graph and $\mathbf{M}$ the corresponding $V \times V$ adjacency matrix. Prove that G is a connected graph if and only if the second eigenvalue λ_2 of $\mathbf{M}$ is strictly less than d.

[Comment: Thus the spectrum of G's adjacency matrix $\mathbf{M}$ "knows" whether it's connected or not!]

Exercise 28

Continuing the previous exercise, let $G = (V, E)$ be a connected d-regular undirected graph and $\mathbf{M}$ the corresponding $V \times V$ adjacency matrix. Prove that G is bipartite if and only if $\lambda_n = -d$.

Lecture 10 Exercises

Exercise 29

Prove that separation between eigenvalues is in general necessary for eigenvectors to be stable under small perturbations (as in the Davis-Kahan theorem). Specifically, show that for every $\epsilon > 0$, there is a matrix $\mathbf{M}$ and a perturbation matrix $\mathbf{P}$ with $\|\mathbf{P}\| \leq \epsilon$ such that (for some i) $\mathbf{M}$ and $\mathbf{M} + \mathbf{P}$ have orthogonal ith eigenvectors.

Exercise 30

Consider the planted clique problem with clique size k at least $c\sqrt{n \log n}$, where n is the number of vertices and c is a sufficiently large constant. Prove that, with high probability (i.e., going to 1 as $n \to \infty$), the vertices of the planted clique are the same as the k vertices with the largest degrees.

[Hint: look up Chebyshev's inequality in the large deviation notes posted to the course Web site. In $\mathcal{G}(n, \frac{1}{2})$, what is the standard deviation of the degree of a vertex?]

Exercise 31

Recall the "expected adjacency matrix" $\widehat{\mathbf{M}}$ for the planted clique problem:

$$\hat{\widehat{\mathbf{M}}} = \left[\begin{array}{ccc|ccc} 1 & \cdots & 1 & \frac{1}{2} & \cdots & \frac{1}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 1 & \frac{1}{2} & \cdots & \frac{1}{2} \\ \hline \frac{1}{2} & \cdots & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{2} & \cdots & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \end{array}\right]$$

Given a closed-form expression for the first eigenvector and eigenvalue of $\widehat{\mathbf{M}}$, as a function of the number of vertices n and the planted clique size k.

Exercise 32

In this exercise we'll finish the proof of correctness of our spectral algorithm for planted clique, by showing that the final clean-up step recovers the planted clique with high probability.

Formally, consider an instance $G = (V, E)$ of planted clique, with planted clique C^* of size $k = \Theta(\sqrt{n})$. Let $C(G)$ denote the set C computed by the canonical spectral algorithm (i.e., by the first three steps) given the graph G. Let $\mathcal{E}$ denote the event that $C(G)$ contains at least $\frac{5}{6}k$ vertices of C^*. (We proved in lecture that $\mathbf{Pr}[\mathcal{E}] = 1 - o(1)$.) Prove that, conditioned on $\mathcal{E}$, with high probability over G, the following holds: a vertex $i \in V$ has at least $\frac{5}{6}k - 1$ neighbors in $C(G)$ if and only if $i \in C^*$.

[Hints: By symmetry, it is enough to consider the case where C^* is deterministic, say $C^* = \{1, 2, \ldots, k\}$. Chernoff bounds will be useful. But be careful completing the argument: G and $C(G)$ are dependent random variables.]

Exercise 33

Prove that there is an algorithm running in $n^{O(\log n)}$ time[1] that correctly computes a maximum clique with high probability (approaching 1 as $n \to \infty$) over the distribution of instances in the planted clique problem. (Here n is the number of vertices, and the planted clique size k can be anything in $\{1, 2, \ldots, n\}$.)

[Hint: You can assume without proof that the size of the maximum clique in a random graph from $\mathcal{G}(n, \frac{1}{2})$ is at most $3 \log_2 n$ with high probability. For $k \geq 10 \log_2 n$, say, suppose you correctly guessed $10 \log_2 n$ vertices of the planted clique C^*? How would you figure out the rest of them?]

Problems

Problem 16

(15 points) In this problem we study the planted bisection problem in the relatively undemanding regime where there is a constant gap between the edge densities p and q. Specifically, assume that $p \geq c_1$ and $q \leq p - c_2$ for constants $c_1, c_2 > 0$. Consider the following simple combinatorial algorithm (given an instance $G = (V, E)$).

A Combinatorial Algorithm for Planted Bisection

1. Choose a vertex $v \in V$ arbitrarily.
2. Let A denote the $\frac{n}{2}$ vertices that have the fewest common neighbors with v.
3. Let B denote the rest of the vertices (including v) and return (A, B).

Prove that, with high probability over the random choice of G (approaching 1 as $n \to \infty$), this algorithm exactly recovers the planted bisection.[2]

[Hint: compute the expected number of common neighbors for pairs of vertices on the same and on different sides of the planted partition. Use the Chernoff bound.]

[1] Also known as "quasi-polynomial-time."

[2] Formally, exact recovery means that the algorithm outputs a bisection (A, B) such that either $A = A^*$ and $B = B^*$ or else $A = B^*$ and $A^* = B$, where (A^*, B^*) is the planted partition.

Problem 17

In Lectures #9–10 we proved that, in n-vertex planted bisection instances with $p = \Omega(1)$ and $p - q = \Omega(\sqrt{\frac{\log n}{n}})$, the canonical spectral algorithm returns a cut (A, B) that misclassifies at most $\frac{n}{32}$ vertices (with high probability). That is, we can label the sides of the planted bisection (A^*, B^*) so that the number of vertices of A^* in B plus the number of vertices of B^* in A is at most $\frac{n}{32}$. Getting most but not all of the planted solution is known as *partial recovery.* This problem describes how to modify the canonical spectral algorithm to get exact recovery of (A^*, B^*), with high probability.

Here is the new (randomized) algorithm (given instance $G = (V, E)$):

The Final Planted Bisection Algorithm

1. Randomly partition the vertex set V into two equal-size groups, V_1 and V_2. Let H_1 and H_2 denote the subgraphs of G induced by V_1 and V_2.
2. Run the canonical spectral algorithm separately on H_1 and H_2, to obtain cuts (A_1, B_1) of H_1 and (A_2, B_2) of H_2.
3. Place each vertex $v \in V_1$ into either $\hat{A}_1$ or $\hat{B}_1$, according to whether v has more neighbors in A_2 or B_2 (breaking ties arbitrarily). Similarly, place each vertex $v \in V_2$ into either $\hat{A}_2$ or $\hat{B}_2$, according to whether v has more neighbors in A_1 or B_1.
4. Return either the cut $(\hat{A}_1 \cup \hat{A}_2, \hat{B}_1 \cup \hat{B}_2)$ or the cut $(\hat{A}_1 \cup \hat{B}_2, \hat{B}_1 \cup \hat{A}_2)$, whichever one is a bisection with fewer crossing edges. (If neither is a bisection, the algorithm fails.)

(a) (12 points) Assume for simplicity of analysis that n is a multiple of 4 and that the algorithm chooses V_1, V_2 so that each contains exactly half of A^* and half of B^*.[3]

Under this assumption, prove that when $p \geq c_1$ and $p - q \geq c_2\sqrt{\frac{\log n}{n}}$ for sufficiently large constants $c_1, c_2 > 0$, the algorithm above exactly recovers the planted bisection (with probability approaching 1 as $n \to \infty$).

[Hint: the "principle of deferred decisions" is your friend. You may also need Chebyshev's inequality.]

(b) (3 points) Why bother with the first and fourth steps? Why not just run the canonical spectral algorithm on the entire graph G to get a cut (A, B), and return the cut $(\hat{A}, \hat{B})$, where $\hat{A}$ (resp., $\hat{B}$) is the set of vertices that have more neighbors in A (resp., B) than B (resp., A)?

Problem 18

This problem fills in the gaps of the proof sketch from lecture that the operator norm $\|\mathbf{R}\|$ of the perturbation matrix $\mathbf{R}$ is $O(\sqrt{n \log n})$. For this problem, assume that $\mathbf{R}$ is a random symmetric matrix, where each R_{ij} lies in $[-1, 1]$ (with probability 1) and $\mathbf{E}[R_{ij}] = 0$. Also, all of the entries R_{ij} with $i \leq j$ are independent from one another. (The entries R_{ij} with $i > j$ have their values forced by the symmetry constraint.)

For part (a) below, you should assume and use the following version of *Hoeffding's inequality*, which acts like a Chernoff bound but for additive error.[4] Let $X_1, \ldots, X_n$ be independent random variables, where X_i takes on values in $[a_i, b_i]$. Let $S = X_1 + \cdots + X_n$ denote the sum and $\mu = \mathbf{E}[S]$ its expectation. Then for every $t \geq 0$,

$$\mathbf{Pr}[|S - \mu| \geq t] \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

[3] A straightforward Chernoff bound argument shows that this will be "almost true" with high probability, and so H_1 and H_2 each contain a planted cut with approximately balanced sides. With minor additional work, the partial recovery guarantee of the canonical spectral algorithm can be extended from planted bisections to planted approximately-balanced cuts. Further details are left to the reader.

[4] The proofs of the two bounds are similar, see CS265.

(a) (5 points) Prove that for every unit vector $\mathbf{u}$ and $t \geq 0$,

$$\mathbf{Pr}\left[|\mathbf{u}^T\mathbf{R}\mathbf{u}| \geq t\right] \leq 2e^{-t^2/2}.$$

[Hint: note that for a unit vector $\mathbf{u}$, $1 = \|\mathbf{u}\|^2 \cdot \|\mathbf{u}\|^2 = \sum_{i=1}^n u_i^4 + 2\sum_{i<j} u_i^2 u_j^2$.]

(b) (5 points) An ϵ-*net* of a metric space (X, d) is a set $N \subseteq X$ such that, for every $x \in X$, there exists $y \in N$ such that $d(x, y) < \epsilon$.

Prove that there is a constant $c > 0$ such that, for all $n \geq 1$ and $\epsilon \in (0, 1)$, the unit ball $\{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1\}$ in n-dimensional Euclidean space has an ϵ-net of size $(\frac{c}{\epsilon})^n$.

[Hints: use a greedy algorithm to construct the net and a volume argument to prove termination. You don't need to explicitly compute any volumes—just assume and use the fact that the (n-dimensional) volume of a δ ball centered at a point $\mathbf{u}$ (i.e., the set $\{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{u} - \mathbf{v}\| \leq \delta\}$) is δ^n times the volume of a unit ball.]

(c) (5 points) Let N be an ϵ-net of the unit sphere in $\mathbb{R}^n$. Prove that there is a constant $c > 0$ such that the following holds, with probability 1 over the choice of $\mathbf{R}$: if $|\mathbf{u}^T\mathbf{R}\mathbf{u}| \leq t$ for some $t \leq n$ and every $\mathbf{u} \in N$, then $|\mathbf{u}^T\mathbf{R}\mathbf{u}| \leq t + c \cdot \epsilon n$ for every unit vector $\mathbf{u}$.

[Hint: what's a trivial upper bound on $\|\mathbf{R}\|$ and hence $\mathbf{v}^T\mathbf{R}\mathbf{v}$?]

(d) (5 points) Conclude that with high probability (approaching 1 as $n \to \infty$), $\|R\| = O(\sqrt{n \log n})$.[5]

Problem 19

Continuing Problem 18, in this problem we prove a tighter bound of $O(\sqrt{n})$ on the size $\|\mathbf{R}\|$ of the perturbation matrix $\mathbf{R}$ in that problem (with high probability). This sharper bound is not relevant for our planted bisection recovery result in Lecture #9, but it is important for our planted clique recovery result in Lecture #10.[6]

Throughout this problem, p^* denotes the probability that, for a fixed unit vector $\mathbf{u}$, a random unit vector $\mathbf{v}$ is close to $\mathbf{u}$, in the sense that $\langle \mathbf{u}, \mathbf{v} \rangle \geq \sqrt{3}/2$. (By symmetry, p^* does not depend on $\mathbf{u}$.)

(a) (5 points) Fix an arbitrary symmetric matrix $\mathbf{A}$. Prove that for a uniformly random unit vector $\mathbf{v}$,

$$\mathbf{Pr}\left[|\mathbf{v}^T\mathbf{A}\mathbf{v}| \geq \frac{1}{2}\|A\|\right] \geq p^*.$$

[Hint: take $\mathbf{u}$ equal to the eigenvector corresponding to the eigenvalue of $\mathbf{A}$ with the largest magnitude.]

(b) (4 points) For arbitrary $t \geq 0$, consider the expression

$$\mathbf{Pr}\left[\|R\| \geq t \text{ and } |\mathbf{v}^T\mathbf{R}\mathbf{v}| \geq \frac{1}{2}\|R\|\right], \tag{1}$$

where the probability is over both the choice of $\mathbf{R}$ and the choice of a uniformly random unit vector $\mathbf{v}$. Prove that the expression in (1) is at least $p^* \cdot \mathbf{Pr}[\|R\| \geq t]$.

(c) (4 points) Prove that the expression in (1) is at most $2e^{-t^2/8}$.

[Hint: Use Problem 18(a).]

[5]In part (b) you only exhibited a small ϵ-net for the unit ball, rather than for the unit sphere, which is not quite the same thing. In this part you can assume the easy-to-believe fact that the sphere also admits an ϵ-net of the same size (which can be proved, for example, by replacing volumes in the argument in (b) with surface areas).

[6]With considerably more work, it's even possible to improve the bound to $\sqrt{n}$ plus lower-order terms (i.e., the leading coefficient is 1).

(d) (9 points) Prove that

$$p^* \geq \frac{c_1}{n2^n},$$

where c_1 is a sufficiently small constant (independent of n).

[Hints: you can use without proof any of the formulas on the Wikipedia page for "Volume of an n-ball." Note that the set of unit vectors $\mathbf{v}$ for which $\langle \mathbf{u}, \mathbf{v} \rangle \geq \sqrt{3}/2$ is a spherical cap (i.e., the intersection of a sphere and a halfspace). The boundary of this spherical cap itself defines a sphere, in one lower dimension, with radius $\frac{1}{2}$ (why?). (Draw pictures of this in two and three dimensions!) Lower bound the surface area of the spherical cap by the surface area of the induced lower-dimensional sphere.]

(e) (3 points) Conclude that $\mathbf{Pr}[\|R\| \geq t] \leq \frac{1}{n}$ provided $t \geq c_2\sqrt{n}$, where c_2 is a sufficiently large constant.

Problem 20

(15 points) In Lecture #10 we showed that a planted clique of size $k = c_1\sqrt{n}$ can be recovered exactly in polynomial time (with high probability), provided c_1 is a sufficiently large constant. Piggyback on this result to prove that, for every constant $\epsilon > 0$, planted cliques of size $\epsilon\sqrt{n}$ can be recovered exactly in polynomial time (with high probability). The running time of your algorithm is allowed to depend exponentially on $\frac{1}{\epsilon}$.

[Hints: Enumerate over all subsets of s vertices (where s depends only on c_1 and ϵ). For each such subset S, consider the subgraph G_S induced by the vertices that are adjacent to all of the vertices of S, and run the original recovery algorithm on G_S.]

Problem 21

In this problem we prove a bound relating the eigenvalues of the sum $\mathbf{A} + \mathbf{B}$ of two matrices with the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$. We used this bound in the proof of the David-Kahan theorem (see Lecture #10 notes).

(a) (5 points) Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. Consider the following game:

(i) You pick a k-dimensional subspace U of $\mathbb{R}^n$.

(ii) Your opponent picks a unit vector $\mathbf{v} \in U$.

Your payoff in the game is defined as $\mathbf{v}^T\mathbf{A}\mathbf{v}$. (You want to maximize this, your opponent wants to minimize this.) Assuming that your opponent plays optimally, prove that the maximum payoff you can obtain in this game is precisely the kth largest eigenvalue λ_k of $\mathbf{A}$.

(b) (5 points) Consider an alternative game, with payoffs defined as before:

(i) Your opponent picks a $(n - k + 1)$-dimensional subspace U.

(ii) You pick a unit vector $\mathbf{v} \in U$.

Prove that, again, the maximum payoff you can obtain in this game (assuming optimal play by your opponent) is precisely λ_k.

(c) (5 points) Let $\mathbf{A}, \mathbf{B}$ be $n \times n$ symmetric matrices with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and $\mu_1 \geq \cdots \geq \mu_n$. Prove that, for $i = 1, 2, \ldots, n$, the ith eigenvalue ν_i of $\mathbf{A} + \mathbf{B}$ satisfies

$$\nu_i \leq \lambda_i + \underbrace{\max_{j=1}^{n} |\mu_j|}_{:=\|\mathbf{B}\|}.$$

[Hint: Use (b), and the fact that $\mathbf{v}^T(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{v}^T\mathbf{A}\mathbf{v} + \mathbf{v}^T\mathbf{B}\mathbf{v}$.]