Graphs and Networks	**out:** November 29, 2007

Problem Set 5

Lecturer: Daniel A. Spielman	**due:** December 13, 2007

1 Introduction

There is 1 section to this problem set. It contains the 3 problems for everyone. Some of them are hard, so just do as much as you can. If you are doing a final project, you can skip this problem set.

2 Homework Policy

You may discuss the problems with other students. But, you must write your solutions independently, drawing on your own understanding. You should cite any sources that you use on the problem sets other than the textbook, TA and instructor. This means that you should list your collaborators.

You **may not** search the web for solutions to similar problems given out in other classes. If you think this policy needs any clarification, please let me know.

3 Corrections

- In problems 2a and 2b, the correct expressions were $c_1 n r^2$ and $c_2 n (\log n) r^2$, not $/r^2$ as previously stated.
- In problem 2b, you are asked to bound the maximum degree in the graph.
- In problem 2d, when I write $\mathbf{Var}[x]$, I mean to set $\mu = \frac{1}{n}\sum x_i$, and $\mathbf{Var}[x] = \frac{1}{n}\left(\sum_i (x_i - \mu)^2\right)$.

4 Problems

1. **Percolation in Random Geometric Graphs**

   Given a set of points $P \subset \mathbb{R}^2$ and a radius r we form a graph with vertex set P and edge set

   $$\{(x, y) : \|x - y\| \leq r\}.$$

   This is equivalent to centering a disk of radius $r/2$ at every point and drawing edges between points whose disks touch. In this problem, we will show that if P is chosen at random according to a Poisson process with unit density, and if r is sufficiently large, then the resulting graph contains an infinite connected component with non-zero probability (actually, the probability is 1).

Recall that under this distribution the probability that there are exactly k points in any region of area A is

$$e^{-A}A^k/k!,$$

and that events in disjoint areas are independent.

(a) Couple the process to bond-percolation on the regular grid of side-length $r/2$. That is, consider the grid graph with vertices $(ir/2, jr/2)$ for i, j integers. Find a way of associating to each edge e of the grid graph a region R_e of $\mathbb{R}^2$ so that if there is a point of $p \in P$ in this region R_e, then the disk of radius $r/2$ centered at p contains both endpoints of the grid edge e. Make sure that the regions corresponding to different edges are disjoint.

(b) Prove that there is a constant value r_0 so that if $r > r_0$, then the graph contains an infinite connected component with non-zero probability. You may use the fact that in the regular grid, there is an infinite component with non-zero probability if each edge appears (is open) with probability greater than 1/2.

2. **Laplacian Eigenvalues of Geometric Graphs** In this problem, we will again consider geometric graphs. However, in this case they will be finite. We will prove that second-smallest Laplacian eigenvalue of a geometric graph with n vertices and maximum degree d_{max} is at most cd_{max}^2/n, for some constant c.

Given a set of points $\{p_1, \ldots, p_n\} \subseteq \mathbb{R}^2$ and a radius r, we define the graph $G(P, r)$ to be the graph with vertex set $\{1, \ldots, n\}$ and edge set

$$E = \{(i, j) : \mathrm{dist}(p_i, p_j) \leq r\}.$$

For simplicity, we will assume in this problem that the points P lie in the unit square: $P \subseteq [0, 1]^2$. First, let's get a handle on how the degrees of vertices behave in random instances.

(a) Prove that if P consists of n randomly chosen points in $[0, 1]^2$, then there exists a constant c_1 so that the expected degree of each vertex of $G(P, r)$ is at most c_1nr^2.

(b) For $r \geq 1/\sqrt{n}$, prove that if P consists of n randomly chosen points in $[0, 1]^2$, then there exists a constant c_2 so that the maximum degree over the vertices of $G(P, r)$ is at most $c_2n(\log n)r^2$, with probability at least $1 - 1/n$.

For the next parts, let $P = \{p_1, \ldots, p_n\}$ and let $p_i = (x_i, y_i)$. Let d_{max} be an upper bound on the maximum degree of $G(P, r)$. We will prove an upper bound on the second-smallest eigenvalue, γ_2, of the Laplacian of the graph (see the first two displayed equations on page 3 of my notes from Lecture 7) by using one of the vectors x or y.

(c) Prove that $x^TLx \leq d_{max}nr^2$.

(d) We cannot just use a vector such as x or y to upper bound γ_2, because they will probably not be orthogonal to the all-1s vector. Let $\hat{x}$ be the vector obtained by subtracting off a multiple of the all-1s vector from x so that the result is orthogonal to the all-1s vector. Prove that $\hat{x}^T\hat{x} = n\mathbf{Var}[x]$.

(e) Prove that if for some s, $\mathbf{Var}[x]$ and $\mathbf{Var}[y]$ are less than $s^2/4$, then there exists a square of side length $2s$ containing at least $n/2$ of the points of P. (hint: use Chebyshev's inequality).

(f) Prove that there exists a constant c_3 so that the number of points of P in a square of side $2s$ is at most $c_3 d_{max} s^2/r^2$. (hint: consider the average ply of a point in the box).

(g) Prove that there exists a constant c_4 so that

$$\max\left(\hat{x}^T\hat{x}, \hat{y}^T\hat{y}\right) \geq \frac{c_4 n^2 r^2}{d_{max}}.$$

(hint: use parts (e) and (f))

(h) Prove that there exists a constant c so that

$$\gamma_2 \leq c d_{max}^2/n.$$

(hint: use parts (g) and (c))

3. **Nearest Neighbor Graphs**

Let $P = \{p_1, \ldots, p_n\}$ be a set of points in $\mathbb{R}^2$ for which all pair-wise distances between the points are distinct. The nearest-neighbor graph $NNG(P)$ is the directed graph with vertex set $\{1, \ldots, n\}$ and edge set

$$\{(i,j) : \|p_i - p_j\| < \|p_i - p_k\|, \text{ for all } k \neq j\}.$$

That is, each node has a directed edge pointing to its nearest neighbor. Note that each node only has out-degree 1.

(a) Prove that the assumption that all pair-wise distances are unique implies that the graph will break into components, each of which consists of a 2-cycle and a bunch of nodes that can reach the 2-cycle by a unique directed path.

Let ℓ be the length of the longest directed path in P. One can show that every component of P has at most $c_1\ell^5$ vertices for some constant c_1. Strangely, this result is tight. We will now establish some ingredients needed in a proof of a slightly weaker result: that each component has at most $c_2\ell^9$ vertices, for some constant c_2.

Assume without loss of generality that the graph $NNG(P)$ is connected, that its 2-cycle is on vertices p_1 and p_2, and that the distance between p_1 and p_2 is 1. Let x be the midpoint of the line segement from p_1 to p_2. Let B_i be the ball of radius 8^i around x, and let S_i be the set of points from P inside ball B_i but not in B_{i-1}.

(b) Prove that there exists a constant c_3 so that for all i, $|S_i| \leq c_3\ell^2$. (hint: the first edge on a directed path must be the longest)

(c) Let E_i be the set of edges from vertices outside B_i to vertices inside B_{i-1}. Prove that, for all i, $|E_i| \leq 6$. (If you cannot prove it for 6, just try to prove it for some constant).

Remark: Feel free to substitute any constant other than 8 in the above problem.

These geometric facts are sufficient to establish the $c_2\ell^9$ bound using some combinatorial arguments, but I will not assign that part of the problem.