Problem Set 4-Grad

Lecturer: Daniel A. Spielman **due:** November 28, 2006

1 Corrections Made

1. (11/13/06) Clarified definition of $\mathrm{sp}(S)$ in problems 4 and 5.
2. (11/27/06) Fixed two details in problem 5: replaced $\alpha k/2$ by αk, and corrected the formula for the number of matchings in a graph with $2n$ vertices.

2 Problems

1. **Sparsity and Conductance** For a graph $G = (V, E)$ and a partition of its vertices $C_1, \ldots, C_k$, we define the sparsity of the partition by

$$\mathrm{sp}(C_1, \ldots, C_k) = \sum_{i=1}^{k} \frac{|\partial(C_i)|}{|C_i|},$$

where $\partial(C_i)$ denotes the set of edges leaving C_i. Define the conductance of a partition by

$$\phi(C_1, \ldots, C_k) = \sum_{i=1}^{k} \frac{\mathrm{vol}(\partial(C_i))}{\mathrm{vol}(C_i)},$$

where vol() of a set of edges is the sum of their weights, and vol() of a set of vertices is the sum of their weighted degrees.

Describe a graph in which the partition C_1, C_2 minimizing $\mathrm{sp}(C_1, C_2)$ and the partition D_1, D_2 minimizing $\phi(D_1, D_2)$ satisfy $C_1 \cap D_1 = \emptyset$ and $|C_2 \cap D_2| > n/4$, where n is the number of vertices in the graph.

If you cannot prove that the obviously minimizing partitions are minimizing, just make your intuitive argument as precise as possible.

2. **Assortativity** Given a graph $G = (V, E)$, we define its assortativity as follows. Let p_k be the fraction of vertices of degree k. Let

$$q_k = \frac{kp_k}{\sum_j jp_j}.$$

The quantity q_k is the probability that a random endpoint of a randomly chosen edge has degree k. (If you look at Newman's paper, he defines this as q_{k+1}, but I think he's got it wrong. In particular, his equation (4) agrees with my interpretation) Define $e_{j,k}$ to be the probability that a randomly chosen edge has degree j at its first endpoint and degree k at the other. (that is, we choose a random edge and a random endpoint)

Define

$$\sigma_q^2 \stackrel{\text{def}}{=} \sum_k k^2 q_k - \left(\sum_k k q_k \right)^2 .$$

Finally, define

$$r(G) = \frac{1}{\sigma_q^2} \sum_{j,k} jk(e_{j,k} - q_j q_k).$$

a. find a graph G with $r(G) = 1$.

b. find a graph G with $r(G) = -1$.

3. Let G be a random directed graph on n vertices in which each vertex has out-degree 1. That is, for every vertex i, we choose a j at random and add edge (i, j) to the graph. For a vertex v, let $R(v)$ denote the set of vertices reachable from v.

   a. Prove that for a fixed vertex v, $\mathrm{P}\left[|R(v)| < \sqrt{n}/10\right] < 1/3$.

   b. Prove that for a fixed vertex v, $\mathrm{P}\left[|R(v)| > 10\sqrt{n}\right] < 1/3$.

   Note that these constants are pretty loose. Much tighter bounds are possible.

4. **You can't cluster a random graph.** Consider a random graph distributed according to $G(n, p)$, with $p = 10 \ln n/(n - 1)$. Prove that there exists an absolute constant $\alpha > 0$ such that

$$\mathrm{P}\left[\min_S \mathrm{sp}(S) < \alpha\right] < 1/3,$$

   where

$$\mathrm{sp}(S) = \frac{|\partial(S)|}{\min(|S|, |V - S|)}.$$

   Note that you can also prove this for $\phi(S)$, it's only slightly trickier.

5. **You can't cluster a cycle plus a matching.** Let G be a cycle plus a random matching. Prove that there exists an absolute constant $\alpha > 0$ such that

$$\mathrm{P}\left[\min_S \mathrm{sp}(S) < \alpha\right] \to 0,$$

   as n goes to infinity, where the minimum is taken over all sets of at most half the vertices.

   a. First, prove that it is unlikely that any particular set of $k \leq n/2$ vertices has fewer than αk matching edges leaving it.
   Hint: There are many ways to do this. One way involves observing that the number of matchings on a graph with $2n$ vertices is $(2n - 1)!! \stackrel{\text{def}}{=} (2n - 1)(2n - 3)(2n - 5) \cdots (1)$. However you do it, you might find useful the bound

$$n! \geq n^n/e^n.$$

b. Now, if you had to sum over all sets of size k, there is no way you could prove the claimed result. But, you don't have to! It is only necessary to consider sets S with fewer than αk cycle edges leaving. Prove an upper bound on the number of such sets, and then prove the claimed result.