princeton university Spr'12 cos 521: Advanced Algorithms

Homework 3

Out: *Mar 16* Due: *Mar 30*

You can collaborate with your classmates, but be sure to list your collaborators with your answer. If you get help from a published source (book, paper etc.), cite that. Also, limit your answers to one page or less —you just need to give enough detail to convince the grader. If you suspect a problem is open, just say so and give reasons for your suspicion.

§1 Modify the approximation scheme for Euclidean TSP given in class to get an approximation scheme for k-MST. (k-MST is the problem of computing the set of k points whose minimum spanning tree is the smallest. Here $k \leq n$is part of the input.)

§2 The *minimum vertex cover* of a graph is a subset of its vertices containing at least one endpoint of every edge. Give an algorithm that finds the minimum vertex cover of a graph with n vertices in time at most 2^{cn} for some constant $c < 0.9$.

§3 Let $X_1, \ldots, X_n$ be independent geometric random variables with parameter $p = 1/2$ (this is the number of fair coin flips before seeing a head; $\mathbf{P}(X_i = t) = 1/2^t$). Show that

$$\mathbf{P}\left(\sum_{i \leq n} X_i > (1+\epsilon)2n\right) \leq \left(\frac{1+2\epsilon}{1+\epsilon}\right)^{-2(1+\epsilon)n} (1+2\epsilon)^n$$

and simplify this to an exponential tail bound.

§4 Consider a random graph G on n vertices in which every edge is included independently with probability $p = 10 \log n/n$. Show that G is connected with high probability. Does this continue to hold when $p = \log n/n$? What about $p = 10/n$?

Consider a random *bipartite* graph on vertex set $L \cup R$ where $|L| = |R| = n$ where each edge is included with probability $10/n$. Show that there is a constant c for which every $S \subset L$ of size $|S| \geq cn$ has at least $\Omega(n)$ edges leaving it, with high probability.

§5 The Johnson-Lindenstrauss lemma tells us that in order to preserve pairwise distances between n vectors up to a $1 \pm \epsilon$ factor, it is sufficient to project onto $k = O(\log n/\epsilon^2)$ dimensions. What is the required target dimension if we only want to preserve 99% of the pairwise distances?

§6 Let $g_1, \ldots, g_n$ be standard Gaussian random vectors in $\mathbf{R}^n$ (i.e., each coordinate is a standard univariate Gaussian). Show that they are almost orthogonal to each other with high probability and obtain a bound on the maximum inner product between any pair. What happens if you take n^2 vectors?

§7 One way to compute the top eigenvector of a symmetric matrix A is to choose a random vector r and take powers $Ar, A^2r, \ldots A^tr$. Assume A has distinct eigenvalues

$\lambda_1 > \lambda_2 > \lambda_3, \dots \lambda_n$. How large must t be to obtain a vector $v_t := A^t r$ with

$$\frac{\langle v_t, u_1 \rangle}{\|v_t\| \|u_1\|} \geq 1 - \epsilon,$$

where u_1 is the top eigenvector? Describe a similar method for computing the second eigenvector u_2.

§8 Compute the spectrum of the walk matrix $W = \frac{1}{2}A$ of the cycle on n vertices. (Hint: The eigenvectors are sines and cosines of different frequencies. To ease calculations, consider the mappings $x \mapsto e^{i\theta x}$ for various θ.)

§9 The $n-$dimensional hypercube has the set of all $n-$bit strings $\{0,1\}^n$ as its vertices, with edges between strings at Hamming distance 1. Compute the its eigenvalues and eigenvectors, with multiplicities. (Hint: Consider the parity functions $\chi_S(x) = (-1)^{\langle x, \mathbf{1}_S \rangle}$ for subsets $S \subset [n]$.)

§10 Show that the mixing time of a connected regular undirected unweighted graph on n vertices is always bounded by a polynomial in n. Give an example of a directed graph for which the random walk started from some point takes $2^{\Omega(n)}$ time to mix.