Spectral Graph Theory and its Applications Lecture 16

Random Graphs, II

Lecturer: Daniel A. Spielman October 28, 2004

16.1 The Problem Set

Problem 2c is false. It is hereby removed from the problem set.

16.2 Wigner's Semi-circle Law

Let R be a random symmetric matrix with independent uniform ± 1 entries. Let $W_n(x)$ denote the expected number of eigenvalues of such an n-by-n matrix that are less than x. Then,

$$\lim_{n \to \infty} W(x 2\sqrt{n}) = W(x),$$

where $W(x) = 0$ for $x \leq -1$, $W(x) = 1$, for $x \geq 1$, and

$$W(x) = \frac{2}{\pi} \int_{-1}^{x} \sqrt{1 - x^2} dx,$$

for $-1 < x < 1$. That is, a histogram of the eigenvalues will look like a semicircle.

16.3 An upper bound

We will prove that for every even k,

$$\mathbf{E}\left[\lambda_{max}(R)^k\right] \leq n(2nk)^{k/2}.$$

By Markov's inequality, this imples that for all $\alpha > 1$,

$$\mathrm{P}\left[\lambda_{max}(R)^k \geq n(2nk)^{k/2} c^k\right] \leq c^{-k}.$$

Taking kth roots, this gives

$$\mathrm{P}\left[\lambda_{max}(R) \geq c(n)^{1/k} \sqrt{2nk}\right] \leq c^{-k}.$$

If we now put $k = \log_2 n$, then $(n)^{1/k} = 2$, so we find

$$\mathrm{P}\left[\lambda_{max}(R) \geq c 2\sqrt{2} \sqrt{n \log n}\right] \leq c^{-\log_2 n}.$$

If we'd proved a stronger bound on the expectation, we wouldn't have the $\log n$ term floating around. Today, we'll prove the bound claimed, and give some indication as to how it can be improved.

16.4 Expectation of Trace

I made a few mistakes on this in class. But, I think the following argument is correct.

Rather than working with $\lambda_{max}(R^k)$ directly, we will prove an upper bound on $\mathbf{E}\left[\mathrm{Tr}\left(R^k\right)\right]$, and observe that for even k

$$\lambda_{max}(R^k) \leq \mathrm{Tr}\left(R^k\right).$$

If the largest eigenvalue is isolated, and k is large, then this bound isn't very far off.

Recall that the trace is the sum of the diagonal entries in a matrix. Also note that

$$R^k(v_0, v_0) = \sum_{v_1,\ldots,v_{k-1}} R(v_0, v_{k-1}) \prod_{i=0}^{k-1} R(v_i, v_{i+1}),$$

and so

$$\mathbf{E}\left[R^k(v_0, v_0)\right] = \sum_{v_1,\ldots,v_{k-1}} \mathbf{E}\left[R(v_0, v_{k-1}) \prod_{i=0}^{k-1} R(v_i, v_{i+1})\right].$$

To simplify this expression, we will recall that if X and Y are independent random variables, then $E(XY) = E(X)E(Y)$. So, to the extent that the terms in this product are independent, we can distribute this expectation accross this product. As the entries of R are independent, up to the symmetry condition, the only terms that are dependent are those that are identical. So, if $\{u_j, w_j\}_j$ is the set of edges that occur in

$$\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{k-2}, v_{k-1}\}, \{v_{k-1}, v_0\}, \tag{16.1}$$

and edge $\{u_j, w_j\}$ appears d_j times, then

$$\mathbf{E}\left[R(v_0, v_{k-1}) \prod_{i=0}^{k-1} R(v_i, v_{i+1})\right] = \prod_j \mathbf{E}\left[R^{d_j}_{(u_j,w_j)}\right].$$

However,

$$\mathbf{E}\left[R^{d_j}_{(u_j,w_j)}\right]$$

is zero if d_j is odd, and one if d_j is even. So, $\mathbf{E}\left[R^k(v_0, v_0)\right]$ equals the number of sequences $v_1, \ldots, v_{k-1}$ such that each edge in (16.1) appears an even number of times. Our goal now is to prove an upper bound on the number of such sequences. Our approach will be to give a way of reconstructing any such sequence.

Let $v_1, \ldots, v_{k-1}$ be a sequence in which every edge $(v_i, v_{i+1\%k})$ appears at least twice, where I use $\%k$ to mean modulo k. Let p denote the number of distinct edges that appears in the sequence, ignoring order. That is, I treat $(1,2)$ and $(2,1)$ as the same edge. Let S denote the set of indices i in which the edge $(v_i, v_{i+1\%k})$ has not appeared before. Then, we have $p = |S|$, and there are at most $\binom{k}{p}$ choices for S.

Given S, we will now describe the sequence using two maps. The first

$$\tau : (\{0, \ldots, k-1\} - S) \to S$$

provides for each time step $i \notin S$ the time step in S in which the edge $\{v_i, v_{i+1}\}$ was first used. The second map,

$$\sigma : S \to \{1, \ldots, n\}$$

gives for each time step $i \in S$ the identity of vertex v_{i+1}.

Let's show that S, τ and σ are enough to reconstruct $v_1, \ldots, v_{k-1}$. We know that the walk starts at node v_0. We will now show that if we know v_i, then we can figure out v_{i+1}. If $i \in S$, then v_{i+1} is just $\sigma(i)$. If $i \notin S$, then we use τ to determine which edge we should traverse. One of its endpoints will be v_i, and the other will be v_{i+1}.

Now, let's count how many possabilities there are for S, σ and τ. We'll settle for a crude upper bound. Given p, there are $\binom{k}{p}$ choices for S. Given S, there are at most n^p choices for σ, and at most p^{k-p} choices for τ. So, for even $k \leq n/2$,

$$\begin{aligned}
\mathbf{E}\left[R^k(v_0, v_0)\right] &\leq \sum_{p=1}^{k/2} \binom{k}{p} n^p p^{k-p} \\
&\leq 2\binom{k}{k/2}(nk/2)^{k/2},
\end{aligned}$$

as the terms are super-geometrically increasing in p,

$$\begin{aligned}
&\leq 2^k (nk/2)^{k/2} \\
&\leq (2nk)^{k/2}.
\end{aligned}$$

Now, the trace is the sum of the diagonal elements, so we find for $k \leq n/2$

$$\mathbf{E}\left[\mathrm{Tr}\left(R^k\right)\right] \leq n(2nk)^{k/2},$$

as desired.