Problem Set 5

Sum of Squares Seminar

October 20, 2017, Due October 30, 2017

Graph Matrix Definition

Recall the definition of the graph matrices R_H

Definition 0.1. *Given a graph H with ordered distinguished sets of vertices U, V, we take R_H to be the matrix such that*

$$R_H(A, B) = \sum_{G' : \exists \sigma : V(H) \to V(G) : \sigma(U)=A, \sigma(V)=B, \sigma(H)=G'} \chi_{E(G')}$$

where A, B are ordered sets of vertices, $\chi_E(G) = (-1)^{|E \setminus E(G)|}$, and we require σ to respect the orderings on U, A, V, B.

Remark 0.2. *This definition is the same as the definition $R_H(A, B) = \sum_{\sigma : V(H) \to V(G) : \sigma(U)=A, \sigma(V)=B} \chi_{\sigma(E(H))}$ up to a constant factor. The advantage of this definition is that it avoids counting the same Fourier character multiple times for a given matrix entry. This difference will not matter for this problem set.*

Problem 1: Decomposing Graph Matrices (15 points)

Express each of the following matrices as a linear combination of the matrices R_H

(a) 5 points: $M_{(a_1,a_2),(b_1,b_2)} = 1$ if a_1, a_2, b_1, b_2 are all distinct and there are precisely 3 edges between a_1, a_2, b_1, b_2 and is 0 otherwise.

(b) 5 points: $M_{(a_1,a_2),(b_1,b_2)} = 2$ if $a_1 = b_1$ and $(a_2, b_2) \in E(G)$ and is zero otherwise.

(c) 5 points: $M_{ab} = (|v \in V(G) \setminus \{a, b\} : (a, v) \in E(G)| - \frac{n-2}{2})(|w \in V(G) \setminus \{a, b\} : (b, w) \in E(G)| - \frac{n-2}{2})$ if $a \neq b$ and $M_{aa} = 0$

Solution

(a) $M = \sum_j \sum_{H:U=\{u_1,u_2\},V=\{v_1,v_2\},U\cap V=\emptyset,|E(H)|=j} c_j R_H$ where $c_0 = \frac{20}{64}$, $c_2 = \frac{-4}{64}$, $c_4 = \frac{4}{64}$, $c_6 = -\frac{20}{64}$, and $c_1 = c_3 = c_5 = 0$

(b) Let H_1 be the graph with distinguished sets of vertices $U = \{u_1, u_2\}$, $V = \{v_1, v_2\}$ where $U \cap V = \{u_1\} = \{v_1\}$ and $E(H_1) = \emptyset$. Similarly, let H_2 be the graph with distinguished sets of vertices $U = \{u_1, u_2\}$, $V = \{v_1, v_2\}$ where $U \cap V = \{u_1\} = \{v_1\}$ and $E(H_1) = \{(u_2, v_2)\}$. $M = R_{H_1} + R_{H_2}$

(c) Let H_1 be the graph with distinguished sets of vertices $U = \{u\}$ and $V = \{v\}$, additional vertices $W = \{w_1, w_2\}$, and edges $E(H_1) = \{(u, w_1), (v, w_2)\}$. Let H_2 be the graph with distinguished sets of vertices $U = \{u\}$ and $V = \{v\}$, additional vertices $W = \{w\}$, and edges $E(H_1) = \{(u, w), (v, w)\}$. $M = \frac{1}{4}(R_{H_1} + R_{H_2})$

Problem 2: Norms of Graph Matrices (15 points)

For each matrix M in problem 1, give probabilistic bounds on $||M||$ and on $||M - E[M]||$

Remark 0.3. *Here it is fine to state that the bounds hold with high probability without stating what that probability is. In case you're curious, in the rough norm bounds there is a factor of $polylog(\frac{1}{\epsilon})$ where ϵ is the probability of failure.*

Solutions

By the matrix norm bounds, these matrices have norm $\tilde{O}\left(n^{\frac{\max_{H:c_H \neq 0}\{|V(H)|-s_H\}}{2}}\right)$ where c_H is the coefficient of H in M and s_H is the minimum size of a vertex separator of H.

(a) With high probability, $||M||$ and $||M - E[M]||$ are both $\Theta(n^2)$. Note that for $||M - E[M]||$, the H with $E(H) = \{(u_1, u_2), (v_1, v_2)\}$ has minimum separator size 0. Thus, for this H, $|V(H)| - s_H = 4$

(b) With high probability, $||M||$ is $\tilde{O}(n)$ as $|V(H_1)| - s_{H_1} = 2$. In $M - E[M]$, the coefficient of R_{H_1} is zeroed out. $V(H_2) - s_{H_2} = 1$, so with high probability, $||M - E[M]||$ is $\tilde{O}(\sqrt{n})$

(c) Here $E[M] = 0$ and with high probability, $||M||$ is $\Theta(n^2)$

Problem 3: Analyzing $N_d(I)$ (30 points)

In this problem, we consider the variance of $N_d(I)$, the number of cliques of size d containing a subset of vertices I.

(a) 10 points: If we decompose $N_4(\emptyset)$ (viewed as a 1×1 matrix) as a linear combination of the graph matrices R_H, which R_H appear and what are their coefficients (up to a constant factor)? For your answer, only use H which have no isolated vertices.

(b) 10 points: Let M be the $n \times n$ matrix with entries $M_{aa} = N_4(\{a\})$ and $M_{ab} = 0$ if $a \neq b$. If we decompose M as a linear combination of the graph matrices R_H, which R_H appear and what are their coefficients (up to a constant factor)? For your answer, only use H which have no isolated vertices (except for a).

(c) 10 points: Give a probabilistic bound (up to constant and logarithmic factors) on how much $N_4(\emptyset)$ and $N_4(\{i\})$ may differ from their expected values. Based on your analysis, what is the main source of this variance? What do you think the pattern is for general $N_d(I)$?

Solutions

(a)

$$N_4(\emptyset) = \sum_{V \subseteq [1,n]:|V|=4} \sum_{E:V(E) \subseteq V} \frac{1}{64}\chi_E = \sum_{E} \sum_{V:V(E) \subseteq V} \frac{1}{64}\chi_E$$

where $V(E)$ is the set of endpoints of E. Thus, $N_4(\emptyset)$ can be decomposed as follows (all H here have empty U, V and have no isolated vertices):

1. $N_4(\emptyset)$ has coefficient $\frac{1}{64}\binom{n}{4}$ for the empty H.
2. $N_4(\emptyset)$ has coefficient $\frac{1}{64}\binom{n-2}{2}$ for the H consisting of a single edge.
3. $N_4(\emptyset)$ has coefficient $\frac{1}{64}(n-3)$ for the H consisting of two edges with one common endpoint.
4. $N_4(\emptyset)$ has coefficient $\frac{1}{64}(n-3)$ for the H consisting of a triangle.
5. $N_4(\emptyset)$ has coefficient $\frac{1}{64}$ for all H such $|V(E(H))| = 4$ where $V(E(H))$ is the set of endpoints of edges of H.

(b)

$$N_4(\{a\}) = \sum_{V \subseteq [1,n]:|V|=4} \sum_{E:V(E) \subseteq V} \frac{1}{64}\chi_E = \sum_{E} \sum_{V:V(E) \subseteq V} \frac{1}{64}\chi_E$$

where $V(E)$ is the set of endpoints of E. Thus, M can be decomposed as follows (all H here have $U = V = \{u\}$ and have no isolated vertices except for u):

1. M has coefficient $\frac{1}{64}\binom{n-1}{3}$ for the H with no edges.

2. M has coefficient $\frac{1}{64}\binom{n-2}{2}$ for the H consisting of a single edge, one of whose endpoints is u.
3. M has coefficient $\frac{1}{64}(n-3)$ for all H such that $|V(E(H)) \cup \{u\}| = 3$
4. M has coefficient $\frac{1}{64}$ for all H such that $|V(E(H)) \cup \{u\}| = 4$

(c) Using the matrix norm bounds, with high probability $N_4(\emptyset) - E[N_4(\emptyset)]$ is $\tilde{O}(n^3)$. The main source of this variance is the variance of the number of edges in the input graph G. Similarly, with high probability, $N_4(\{a\}) - E[N_4(\{a\})]$ is $\tilde{O}(n^{2.5})$. The main source of this variance is the variance of the degree of a.

More generally, with high probability $N_d(\emptyset) = (1 \pm \tilde{O}(\frac{1}{n}))2^{-\binom{d}{2}}\binom{n}{d}$ and the main source of this variance is the variance of the number of edges in the input graph G. Conditioned on I being a clique, with high probability $N_d(I) = (1 \pm \tilde{O}(\frac{1}{\sqrt{n}}))2^{\binom{|I|}{2}-\binom{d}{2}}\binom{n-|I|}{d-|I|}$ and the main source of this variance is the variance of the number of vertices which are adjacent to every vertex in I.

Problem 4: Analyzing $E[M']$ (10 points)

Recall that $E[M']$ is the matrix with entries $(E[M'])_{IJ} = 2^{\binom{|I|}{2}+\binom{|J|}{2}-\binom{|I\cap J|}{2}} \frac{\binom{k}{|I\cup J|}}{\binom{n}{|I\cup J|}}$ where $|I| = |J| = \frac{d}{2}$. Further recall the D_i and P_i bases for the Johnson scheme. $(D_i)_{IJ} = 1$ if $|I \cap J| = i$ and is 0 otherwise. $(P_i)_{IJ} = \binom{|I\cap J|}{i}$.
Decompose $E[M']$ in terms of the P_i basis (your answer will be a bit messy) and deduce that $E[M']$ has a high minimal eigenvalue.

Solution

Recall that $D_i = \sum_{j=i}^{\frac{d}{2}} (-1)^{j-i}\binom{j}{i}P_j$. We now have that

$$E[M'] = \sum_{i=0}^{\frac{d}{2}} 2^{\binom{\frac{d}{2}}{2}+\binom{\frac{d}{2}}{2}-\binom{i}{2}} \frac{\binom{k}{d-i}}{\binom{n}{d-i}} D_i = \sum_{i=0}^{\frac{d}{2}}\sum_{j=i}^{\frac{d}{2}} (-1)^{j-i}\binom{j}{i} 2^{\binom{\frac{d}{2}}{2}+\binom{\frac{d}{2}}{2}-\binom{i}{2}} \frac{\binom{k}{d-i}}{\binom{n}{d-i}} P_j$$
$$= \sum_{j=0}^{\frac{d}{2}}\sum_{i=0}^{j} (-1)^{j-i}\binom{j}{i} 2^{\binom{\frac{d}{2}}{2}+\binom{\frac{d}{2}}{2}-\binom{i}{2}} \frac{\binom{k}{d-i}}{\binom{n}{d-i}} P_j$$

Since $n >> k$, for each j the dominant term will be the term where $i = j$ so the coefficient of each P_j will be roughly $2^{\binom{\frac{d}{2}}{2}+\binom{\frac{d}{2}}{2}-\binom{j}{2}} \frac{\binom{k}{d-j}}{\binom{n}{d-j}}$ which is much bigger than 0. In particular, the coefficient of $P_{\frac{d}{2}} = Id$ is $\Theta\left(\frac{k^{\frac{d}{2}}}{n^{\frac{d}{2}}}\right)$

Problem 5: Wigner's Semicircle Law (30 points)

Wigner's semicircle law on the spectrum of ± 1 symmetric random matrices says the following. If M is a symmetric ± 1 random matrix then as n goes to infinity, the proportion of eigenvalues between $x\sqrt{n}$ and $(x+dx)\sqrt{n}$ approaches $\frac{1}{2\pi}\sqrt{4-x^2}$. In this problem, we explore why Wigner's semicircle law holds. Let $C_k = \frac{1}{k+1}\binom{2k}{k}$ be the kth Catalan number.

(a) 15 points: Show that for all $k \geq 1$, $E\left[tr\left(\left(MM^T\right)^k\right)\right] = C_k n^{k+1} \pm O(n^k)$ (hint is available)

(b) 10 points: Show that $\int_{x=-2}^{2} \frac{1}{2\pi} x^{2k}\sqrt{4-x^2}dx = C_k$ (hint is available)

(c) 5 points: To the best of your ability, explain why this implies that Wigner's semicircle law holds. One thing you could do is to assume that the eigenvalues of M divided by $\sqrt{n}$ approach some distribution and then argue that this distribution must be $\frac{1}{2\pi}\sqrt{4-x^2}$.

Solutions

(a)

$$E\left[tr\left(\left(MM^T\right)^k\right)\right] = \sum_{a_1,b_1,\cdots,a_k,b_k} \prod_{i=1}^{k} M_{a_i b_i} M_{a_{i+1} b_i}$$

where $a_{k+1} = a_1$. Consider the terms which have nonzero expected value. For these terms, each M_{ab} must appear an even number of times. We can ignore the terms which have at most k distinct indices as these will contribute $O(n^k)$.

Lemma 0.4. *For the terms which have $k+1$ distinct indices and have nonzero expected value:*

(a) For all i, j, $a_i \neq b_j$

(b) We can draw the constraint graph showing which indices are equal to each other as follows. We place the indices $a_1, b_1, \cdots, a_k, b_k$ on a circle. Whenever $a_{i_2} = a_{i_1}$ and there is no j such that $a_j = a_{i_1} = a_{i_2}$ and $i_1 < j < i_2$ then we draw an edge from a_{i_1} to a_{i_2}. If we draw the constraint graph in this way then there will be no edge crossings.

Proof. The base case $k = 2$ is trivial. For $k > 2$, note that there must be a unique index as otherwise there would be at most k distinct indices. Without loss of generality, assume this index is b_1. If so, then we must have that $a_1 = a_2$ so we can contract these indices together, delete b_1, and apply the inductive hypothesis. □

With this lemma in hand, we have the following bijection between constraint graphs on $a_1, b_1, \cdots, a_k, b_k$ and walks of length $2k$ where we go up or down at each step and never go below height 0.

To go from a walk to a constraint graph, for $j \in [0, k-1]$, label the $(2j)$-th vertex of the walk a_{j+1} and label the $(2j+1)$-th vertex of the walk b_{j+1}. The final step of the walk will be to $a_{k+1} = a_1$. Whenever we take a step down, draw a constraint edge between the endpoint of the step and the previous vertex at that height.

Conversely, to go from a constraint graph to a walk, start at a_1, go through the vertices $b_1, a_2, b_2, \cdots, a_k, b_k$ in order. Whenever a new index is encountered, take a step up. Whenever an index is encountered which has been encountered before, take a step down.

For each constraint graph with $k+1$ distinct indices, there are $n^{k+1} - O(n^k)$ different possibilities. Thus, $E\left[tr\left(\left(MM^T\right)^k\right)\right] = C_k n^{k+1} \pm O(n^k)$

(b) Observe that for all $k \geq 1$,

$$\int (cos(\Theta))^{2k} d\Theta = sin(\Theta)(cos(\Theta))^{2k-1} + (2k-1)\int (sin(\Theta))^2 (cos(\Theta))^{2k-2} d\Theta$$
$$= sin(\Theta)(cos(\Theta))^{2k-1} + (2k-1)\int (cos(\Theta))^{2k-2} d\Theta - (2k-1)\int (cos(\Theta))^{2k} d\Theta$$

Thus,

$$2k\int (cos(\Theta))^{2k} d\Theta = sin(\Theta)(cos(\Theta))^{2k-1} + (2k-1)\int (cos(\Theta))^{2k-2} d\Theta$$

which implies that

$$\int_0^{\pi} (cos(\Theta))^{2k} d\Theta = \frac{2k-1}{2k}\int_0^{\pi} (cos(\Theta))^{2k-2} d\Theta = \pi \prod_{j=1}^{k} \left(\frac{2j-1}{2j}\right)$$

Using this and taking the substitution $x = 2cos(\Theta)$,

$$\begin{aligned}\int_{x=-2}^{2} \frac{1}{2\pi} x^{2k}\sqrt{4-x^2}\,dx &= \frac{2^{2k+1}}{\pi}\int_{\Theta=0}^{\pi} (cos(\Theta))^{2k}(sin(\Theta))^2 d\Theta \\ &= \frac{2^{2k+1}}{\pi}\int_{\Theta=0}^{\pi} (cos(\Theta))^{2k} d\Theta - \frac{2^{2k+1}}{\pi}\int_{\Theta=0}^{\pi} (cos(\Theta))^{2k+2} d\Theta \\ &= 2^{2k+1}\left(1 - \frac{2k+1}{2k+2}\right)\prod_{j=1}^{k}\left(\frac{2j-1}{2j}\right) = \frac{2^{2k}}{(k+1)}\prod_{j=1}^{k}\left(\frac{2j-1}{2j}\right)\end{aligned}$$

Since $\prod_{j=1}^{k} (2j-1) = \frac{(2k)!}{2^k (k!)}$ and $\prod_{j=1}^{k} \frac{1}{2j} = \frac{1}{2^k (k!)}$,

$$\int_{x=-2}^{2} \frac{1}{2\pi} x^{2k}\sqrt{4-x^2}\,dx = \frac{1}{k+1}\binom{2k}{k} = C_k$$

as needed.

(c) Assume that the eigenvalues divided by $\sqrt{n}$ approach some distribution $f(x)$, i.e. the proportion of eigenvalues between $x\sqrt{n}$ and $(x+dx)\sqrt{n}$ approaches $f(x)dx$ as $n \to \infty$. Further assuming that $tr\left((MM^T)^k\right)$ is concentrated around it's expectation, for all $k \geq 0$,

$$\lim_{n\to\infty} \frac{E[tr\left((MM^T)^k\right)]}{n^{k+1}} = \lim_{n\to\infty} \frac{1}{n}\sum_i \left(\frac{\lambda_i}{\sqrt{n}}\right)^{2k} = \int_{x=-\infty}^{\infty} x^{2k} f(x)dx = C_k$$

For odd moments, observe that M and $-M$ are equally likely so we must have that for all $k \geq 0$, $\int_{x=-\infty}^{\infty} x^{2k+1} f(x)dx = 0$.

Thus, for all $k \geq 0$, $\int_{x=-\infty}^{\infty} x^k f(x)dx = \int_{x=-\infty}^{\infty} x^k \frac{1}{2\pi}\sqrt{4-x^2}dx$

Note: A more rigorous explanation may be added in the future.

Hints

5a. Use the following characterization of the Catalan numbers. C_k is the number of ways to take a total of k steps up and k steps down. With this characterization of the Catalan numbers, it is sufficient to find a bijection between such walks and constraint graphs on a cycle of length $2k$ with $k+1$ distinct indices.

5b. Take the substitution $x = 2cos(\Theta)$ and use the fact (which can be shown by integration by parts) that for all $k \geq 1$, $\int_0^\pi (cos(\Theta))^{2k} d\Theta = \frac{2k-1}{2k} \int_0^\pi (cos(\Theta))^{2k-2} d\Theta$