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 $M_{aa} = 0$

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.*

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.

(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 $N_d(I)$?

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

Recall that $E[M']$ is the matrix with entries $(E[M'])_{IJ} = 2^{\binom{|I|}{2}+\binom{|J|}{2}-\binom{|I\cap J|}{2}} \left(\frac{k^{|I\cup J|}}{n^{|I\cup J|}}\right)$ 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.

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 $k \geq 1$, $E\left[tr\left(\left(MM^T\right)^k\right)\right] = C_k n^{k+1} + O(n^k)$ (hint is available)

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

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 $k \geq 1$, $\int_0^\pi (cos(\Theta))^{2k} d\Theta = \frac{2k-1}{2k} \int_0^\pi (cos(\Theta))^{2k-2} d\Theta$