CSE 599: Algorithmic Spectral Graph Theory HW #3

Instructor: James R. Lee

This homework is due in two weeks on Wednesday, June 6th. You may work in small groups of 2 or 3. You should reference any discussions with others, books, lecture notes, or other resources that you use in coming up with a solution. You may slip your paper under my door, put it in my mailbox, or email it to me at `jrl@cs.washington.edu`.

Problems

1. **Lovász-Simonovits analysis.** Let $G = (V, E)$ be a d-regular graph on n vertices (without any self-loops), and for any distribution p and number $\mu \in [0, n]$, define

$$I(p, \mu) = \max_{w: V \to [0,1]} \left\{ \sum_{v \in V} w(v) p(v) : \sum_{v \in V} w(v) = \mu \right\} .$$

   Let $W = \frac{1}{2}(I + \frac{1}{d} A)$ be the lazy random walk matrix, where A is the adjacency matrix of G. (Note that W is a symmetric matrix.)

   Define $\Phi = \min_{|S| \leqslant n/2} \frac{|E(S)|}{d|S|}$ to be the expansion constant of the graph. Your goal is to prove that for any integer $k \leqslant n/2$ and any distribution p, we have the inequality

$$I(Wp, k) \leqslant \frac{1}{2} \left[I(p, k - 2\Phi k) + I(p, k + 2\Phi k) \right] . \qquad (1)$$

   Hint: Let S be a set which maximizes $I(Wp, k)$, i.e. such that $I(Wp, k) = \langle \mathbf{1}_S, Wp \rangle$. You should see how much probability mass can flow into S when applying one step of the walk W to p. There are three kinds of edges along which probability mass will flow: Edges with two distinct endpoints in S, edges with exactly one endpoint in S, and self-loops in S.

   Let $\phi = \phi(S)$ be the expansion of S. You should prove (1) with ϕ in place of Φ, and then argue that it also holds for Φ. Use ϕ to calculate the number of edges of the three types. Edges of the first type will yield the first term in (1), while edges of the second and third type will give the second term. The factor 1/2 will come from the fact that every vertex has a self-loop of weight 1/2 (coming from the laziness of the random walk). This is just a sketch. Provide a formal proof.

2. **The Gaussian free field.** You will prove that, for any graph $G = (V, E)$, there is a Gaussian process $\{g_u\}_{u \in V}$ with

$$\mathbb{E}(g_u - g_v)^2 = R_{\text{eff}}(u, v) \quad \text{for all } u, v \in V . \qquad (2)$$

   Fix $v_0 \in V$. Define $\{g_u\}_{u \in V}$ to be the process with $g_{v_0} = 0$ and density proportional to

$$\exp\left(-\frac{1}{2} \langle \vec{g}, L\vec{g} \rangle \right) = \exp\left(-\frac{1}{2} \sum_{u \sim v} (g_u - g_v)^2 \right) ,$$

   where $\vec{g} = (g_u : u \in V)$. This is called the *Gaussian free field on G* (grounded at v_0). Here, L is the combinatorial Laplacian defined by $Lf(u) = \sum_{v \sim u} (f(u) - f(v))$.

(a) Prove that there is such a Gaussian process. Write the distribution in terms of the pseudoinverse of the Laplacian $L^\dagger$. You may use the fact that if $\{\eta_u : u \in V\}$ is a family of i.i.d. $N(0,1)$ random variables, then their density is proportional to $\exp(-\frac{1}{2}\sum_{u\in V}\eta_u^2)$.

(b) Define $\Gamma(u,v) = \frac{1}{2}\left(R_{\text{eff}}(u,v_0) + R_{\text{eff}}(v,v_0) - R_{\text{eff}}(u,v)\right)$. Let $\tilde{\Gamma}$ and $\tilde{L}$ be the corresponding matrices with the row and column corresponding to v_0 deleted. Argue that if $\tilde{L}\tilde{\Gamma} = I$ is true, then (2) is true.

(c) The *cycle identity for hitting times* states that

$$H(u,v_0) + H(v_0,v) + H(v,u) = H(u,v) + H(v,v_0) + H(v_0,u)\,.$$

(See Lemma 10.10 in `http://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf` for a proof.) Also recall that $R_{\text{eff}}(u,v) = 2|E|\cdot\kappa(u,v)$ where $\kappa(u,v) = H(u,v) + H(v,u)$ is the commute time. Use these two facts to prove that

$$\Gamma(u,v) = \frac{1}{4|E|}\left(H(u,v_0) + H(v_0,v) - H(u,v)\right)\,. \tag{3}$$

(d) Use (3) to prove that for $v_0 \notin \{u,v\}$, we have

$$d_u\Gamma(u,v) = \mathbf{1}_{\{u=v\}} + \sum_{w\sim u}\Gamma(w,v)\,,$$

where d_u is the degree of u.

(e) Finally, use the definition of the Laplacian and the preceding equality to prove that $\tilde{L}\tilde{\Gamma} = I$, completing the proof.

3. **Chaining.** Consider the Gaussian process $\{X_n : n \geqslant 2\}$ given by

$$X_n = \frac{g_n}{\sqrt{\log n}},$$

where $\{g_n\}$ is a sequence of i.i.d. N(0,1) random variables. Use the chaining method to prove that

$$\mathbb{E}\sup_n X_n < \infty\,.$$

For the sake of convenience, I recall here the chaining upper bound. Let $\{X_n : n \in \mathbb{N}\}$ be a Gaussian process equipped with the metric $d(i,j) = \sqrt{\mathbb{E}(X_i - X_j)^2}$. Let $T_0 \subseteq T_1 \subseteq \cdots \subseteq \mathbb{N}$ be a sequence of subsets which satisfy $|T_0| = 1$ and $|T_k| \leqslant 2^{2^k}$ for $k \geqslant 1$. Then,

$$\mathbb{E}\sup_{n\in\mathbb{N}} X_n \leqslant O(1)\sup_{n\in\mathbb{N}}\sum_{k\geqslant 0} 2^{k/2} d(n,T_k)\,,$$

where $d(n,T_k) = \min_{m\in T_k} d(n,m)$.