1 Matrix Chernoff

Löwner ordering: We write $A \preceq 0$ to denote that $A$ is PSD. We extend it to a partial ordering between symmetric matrices, defining $A \succeq B$ if $A - B \succeq 0$.

Observe that the spectral sparsification condition (2) from last time can be written as

$$(1 - \varepsilon)L_G \preceq L_G' \preceq (1 + \varepsilon)L_G.$$ 

Matrix Chernoff bound [Tropp, 2012]: Let $X_1, \ldots, X_k$ be independent random $n \times n$ symmetric matrices. Suppose that

$$\forall i \in [k], \quad 0 \preceq X_i \preceq I \quad \text{and} \quad \mu \cdot I \preceq \sum_{i=1}^{k} \mathbb{E}[X_i] \preceq \overline{\mu} \cdot I.$$ 

Then for all $\varepsilon \in [0, 1]$,

$$\Pr \left[ \lambda_{\max}(\sum_{i=1}^{k} X_i) \geq (1 + \varepsilon)\overline{\mu} \right] \leq n \cdot e^{-\varepsilon^2 \overline{\mu} / 3},$$

$$\Pr \left[ \lambda_{\min}(\sum_{i=1}^{k} X_i) \leq (1 - \varepsilon)\mu \right] \leq n \cdot e^{-\varepsilon^2 \mu / 2}.$$ 

2 Spectral Sparsifiers (cont’d)

We now continue to analyze Algorithm SS seen last week.

Proof of Lemma 5: Was seen in class using the cyclic property of trace.

There is also a combinatorial explanation for this equality: $w_e R_{\text{eff}}(e)$ can be shown to be exactly the probability that edge $e$ appears in a random spanning tree of $G$, when the probability to sample
any specific tree is proportional to the product of its edge weights. The expected number of edges in such a random tree is just the sum of these edge probabilities, and clearly it is also $n - 1$.

**Connection to importance sampling:** Lemma 7 below shows that $w_{uv} R_{\text{eff}}(u, v)$ for an edge $uv \in E$ is precisely the maximum possible (over all $x$) relative contribution of this edge to $x^T L_G x = \sum_{ij \in E} w_{ij} (x_i - x_j)^2$. Thus, the sampling probability $p_e$ of an edge is proportional to its worst-case relative contribution to $x^T L_G x$. (Why proportionally and not exactly? because the values $w_{uv} R_{\text{eff}}(u, v)$ could sum up to more than 1.)

We could thus apply the importance sampling theorem with $\lambda = n - 1$ for any specific $x \in \mathbb{R}^V$. However, this would not prove Theorem 3, because the importance sampling theorem provides only weak concentration, that is not strong enough to take a union bound over all $x \in \mathbb{R}^V$.

**Lemma 7:**

$$\forall uv \in E, \quad R_{\text{eff}}(u, v) = \max_{x \in \mathbb{R}^V} \frac{(x_u - x_v)^2}{x^T L_G x}.$$

Observe that we can think of $x$ as a vector of potentials $\phi \in \mathbb{R}^V$, and restate the lemma as an analogue of Thomson’s principle (minimizing energy, but now for potentials):

$$\forall uv \in E, \quad R_{\text{eff}}(u, v) = \left[ \min_{\phi_u - \phi_v = 1} \phi^T L_G \phi \right]^{-1}.$$

Proof hint: Consider a minimizer $\phi$. First show that every $\phi_i$ for $i \neq u, v$ is the weighted average of $\phi_j$ over its neighbors $j \in N(i)$. Then use this minimizer $\phi$ to define an electrical flow $f$, and use this flow to express each side, $R_{\text{eff}}(u, v)$ and $\phi^T L_G \phi$.

**Proof of Theorem 3:** Was seen in class. The basic idea is to use the Matrix Chernoff bound, but since it is “built” for scenarios where the expectation is $\mu I$, we need to rotate/change the basis, achieved by multiplying by $L_G^{-1/2}$. More precisely, we define

$$y_{uv} := L_G^{-1/2} z_{uv},$$

and now claim (as an exercise) that

Exer: Show that

$$(1 - \varepsilon)L_G \preceq L_G' = \sum_{e \in E} w_e' Z_e \preceq (1 + \varepsilon)L_G \quad (1)$$

if and only if (modulo the pseudo-inverse/kernel issue)

$$(1 - \varepsilon)I \preceq L_G^{-1/2} (\sum_{e \in E} w_e' z_e z_e^T) L_G^{-1/2} = \sum_{e \in E} w_e' y_e y_e^T \preceq (1 + \varepsilon)I,$$

where we define

$$y_{uv} := L_G^{-1/2} z_{uv}.$$

Hint: Multiply from left and right by $L_G^{-1/2}$. 

2
Denote the random edge chosen at iteration $i \in [k]$ by $e_i$, and then the random matrix (from above) that we need analyze can be written as

$$M' = \sum_{e \in E} w'_e y_e y_e^\top = \sum_{i=1}^{k} \frac{n - 1}{k \cdot R_{\text{eff}}(e_i)} y_{e_i} y_{e_i}^\top.$$  \hfill (2)

To complete the proof of Theorem 3, apply the matrix Chernoff bound to $\frac{k}{n-1} M' = \sum_{i=1}^{k} \frac{1}{R_{\text{eff}}(e_i)} y_{e_i} y_{e_i}^\top$, (after checking the conditions), and conclude the required bounds on the eigenvalues of $M'$.

**Exer:** Explain how to modify the analysis when the sampling loop in steps 3-5 of Algorithm SS is changed to the following: for each edge $e \in E$, repeat $k' = O(\epsilon^{-2} \log n)$ times, where each repetition increases the weight $w'_e$ (as in step 5) independently with probability $p_e$.

**Exer:** Show how to modify the algorithm and its analysis to use estimates $\tilde{p}_e$ instead of $p_e$ (e.g., maybe these estimates can be computed very quickly), under the assumption that every $\tilde{p}_e \geq p_e$, and that $\sum_{e \in E} \tilde{p}_e \leq C$.

Hint: you may use the preceding exercise.