1 Spectral Sparsifiers

1.1 Graph Laplacians

Let $G = (V, E)$ be an unweighted graph. For notational simplicity, we will think of the vertex set as $V = \{1, \ldots, n\}$. Let $e_i \in \mathbb{R}^n$ be the $i$th standard basis vector, meaning that $e_i$ has a 1 in the $i$th coordinate and 0s in all other coordinates. For an edge $uv \in E$, define the vector $x_{uv}$ and the matrix $X_{uv}$ as follows:

\[ x_{uv} := e_u - e_v \]
\[ X_{uv} := x_{uv}x_{uv}^T \]

In the definition of $x_{uv}$ it does not matter which vertex gets the +1 and which gets the −1 because the matrix $X_{uv}$ is the same either way.

**Definition 1** The Laplacian matrix of $G$ is the matrix

\[ L_G := \sum_{uv \in E} X_{uv} \]

Let us consider an example.

\[
G
\]

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & 0 & -1 & 0 \\
2 & 0 & 0 & 0 \\
3 & -1 & 0 & 1 \\
4 & 0 & 0 & 0 \\
\end{array}
\]

\[
X_{13} =
\]

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & 2 & -1 & 0 \\
2 & -1 & 2 & -1 \\
3 & -1 & -1 & 3 \\
4 & 0 & 0 & -1 \\
\end{array}
\]

-1 indicates an edge from 1 to 3
degree of node 3 = # edges that hit 3
Note that each matrix $X_{uv}$ has only four non-zero entries: we have $X_{uu} = X_{vv} = 1$ and $X_{uv} = X_{vu} = -1$. Consequently, the $u$th diagonal entry of $L_G$ is simply the degree of vertex $u$. Moreover, we have the following fact.

**Fact 2** Let $D$ be the diagonal matrix with $D_{u,u}$ equal to the degree of vertex $u$. Let $A$ be the adjacency matrix of $G$. Then $L_G = D - A$.

If $G$ had weights $w : E \rightarrow \mathbb{R}$ on the edges we could define the weighted Laplacian as follows:

$$L_G = \sum_{uv \in E} w_{uv} \cdot X_{uv}.$$

**Claim 3** Let $G = (V, E)$ be a graph with non-negative weights $w : E \rightarrow \mathbb{R}$. Then the weighted Laplacian $L_G$ is positive semi-definite.

**Proof:** Since $X_{uv} = x_{uv}x_{uv}^T$, it is positive semi-definite. So $L_G$ is a weighted sum of positive semi-definite matrices with non-negative coefficients. Fact 5 in the Notes on Symmetric Matrices implies $L_G$ is positive semi-definite. □

The Laplacian can tell us many interesting things about the graph. For example:

**Claim 4** Let $G = (V, E)$ be a graph with Laplacian $L_G$. For any $U \subseteq V$, let $\chi(U) \in \mathbb{R}^n$ be the characteristic vector of $U$, i.e., the vector with $\chi(U)_v$ equal to 1 if $v \in U$ and equal to 0 otherwise. Then $\chi(U)^T L_G \chi(U) = |\delta(U)|$.

**Proof:** For any edge $uv$ we have $\chi(U)^T X_{uv} \chi(U) = (\chi(U)^T x_{uv})^2$. But $|\chi(U)^T x_{uv}|$ is 1 if exactly one of $u$ or $v$ is in $U$, and otherwise it is 0. So $\chi(U)^T X_{uv} \chi(U) = 1$ if $uv \in \delta(U)$, and otherwise it is 0. Summing over all edges proves the claim. □

Similarly, if $G = (V, E)$ is a graph with edge weights $w : E \rightarrow \mathbb{R}$ and $L_G$ is the weighted Laplacian, then then $\chi(U)^T L_G \chi(U) = w(\delta(U))$.

**Fact 5** If $G$ is connected then $\text{image}(L_G) = \{ x : \sum_i x_i = 0 \}$, which is an $(n-1)$-dimensional subspace.

1.2 Main Theorem

**Theorem 6** Let $G = (V, E)$ be a graph with $n = |V|$. There is a randomized algorithm to compute weights $w : E \rightarrow \mathbb{R}$ such that:

- only $O(n \log n/\epsilon^2)$ of the weights are non-zero, and
- with probability at least $1 - 2/n$,

$$ (1 - \epsilon) \cdot L_G \preceq L_w \preceq (1 + \epsilon) \cdot L_G, $$

where $L_w$ denotes the weighted Laplacian of $G$ with weights $w$. By Fact 4 in Notes on Symmetric Matrices, this is equivalent to

$$ (1 - \epsilon) x^T L_G x \leq x^T L_w x \leq (1 + \epsilon) x^T L_G x \quad \forall x \in \mathbb{R}^n. \quad (1) $$
By (1) and Claim 4, the resulting weights are a graph sparsifier of $G$:

$$(1 - \epsilon) \cdot |\delta(U)| \leq w(\delta(U)) \leq (1 + \epsilon) \cdot |\delta(U)| \quad \forall U \subseteq V.$$ 

The algorithm that proves Theorem 6 is as follows.

- Initially $w = 0$.
- Set $k = 8n \log(n)/\epsilon^2$.
- For every edge $e \in E$ compute $r_e = \text{tr}(X_e L^+ G)$.
- For $i = 1, \ldots, k$
  - Let $e$ be a random edge chosen with probability $r_e/(n - 1)$.
  - Increase $w_e$ by $(n - 1) r_e k$.

**Claim 7** The values $\{r_e/(n - 1) : e \in E\}$ indeed form a probability distribution.

**Proof:** (of Theorem 6). How does the matrix $L_w$ change during the $i$th iteration? The edge $e$ is chosen with probability $r_e/n - 1$ and then $L_w$ increases by $(n - 1) X_e$. Let $Z_i$ be this random change in $L_w$ during the $i$th iteration. So $Z_i$ equals $(n - 1) X_e$ with probability $r_e/n - 1$. The random matrices $Z_1, \ldots, Z_k$ are mutually independent and they all have this same distribution. Note that

$$E[Z_i] = \sum_{e \in E} r_e \cdot \frac{n - 1}{r_e \cdot k} X_e = \frac{1}{k} \sum_e X_e = \frac{L_G}{k}. \quad (2)$$

The final matrix $L_w$ is simply $\sum_{i=1}^k Z_i$. To analyze this final matrix, we will use the Ahlswede-Winter inequality. All that we require is the following claim, which we prove later.

**Claim 8** $Z_i \preceq (n - 1) \cdot E[Z_i]$.

We apply Corollary 2 from the previous lecture with $R = n - 1$, obtaining

$$\Pr[(1 - \epsilon) L_G \preceq L_w \preceq (1 + \epsilon) L_G] = \Pr \left[ (1 - \epsilon) \frac{L_G}{k} \preceq \frac{1}{k} \sum_{i=1}^k Z_i \preceq (1 + \epsilon) \frac{L_G}{k} \right]$$

$$\leq 2n \cdot \exp \left( - \epsilon^2 k / 4(n - 1) \right)$$

$$\leq 2n \cdot \exp \left( - 2 \ln n \right) < 2/n.$$

$\square$

## 2 Appendix: Additional Proofs

**Proof:** (of Claim 7) First we check that the $r_e$ values are non-negative. By the cyclic property of trace

$$\text{tr}(X_e L^+_G) = \text{tr}(x_e^T L^+_G x_e) = x_e^T L^+_G x_e,$$

This is non-negative since $L^+_G \succeq 0$ because $L_G \succeq 0$. Thus $r_e \geq 0$. 


Next, note that
\[ \sum_e \text{tr}(X_e L_G^+) = \text{tr}(\sum_e X_e L_G^+) = \text{tr}(L_G L_G^+) = \text{tr}(I_{\text{im } L_G}), \]
where \( I_{\text{im } L_G} \) is the orthogonal projection onto the image of \( L_G \). The image has dimension \( n - 1 \) by Fact 5, and so
\[ \sum_e r_e = \frac{1}{n-1} \sum_e \text{tr}(X_e L_G^+) = \frac{1}{n-1} \text{tr}(I_{\text{im } L_G}) = 1. \]
\[ \square \]

**Proof:** (of Claim 8). The maximum eigenvalue of a positive semi-definite matrix never exceeds its trace, so
\[ \lambda_{\text{max}}(L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2) \leq \text{tr}(L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2) = r_e. \]
By Fact 8 in the Notes on Symmetric Matrices,
\[ L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2 \preceq r_e \cdot I. \]
So, by Fact 4 in the Notes on Symmetric Matrices, for every vector \( v \),
\[ v^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} \leq v^T v. \] (3)
Now let us write \( v = v_1 + v_2 \) where \( v_1 = I_{\text{im } L_G} v \) is the projection onto the image of \( L_G \) and \( v_2 = I_{\text{ker } L_G} v \) is the projection onto the kernel of \( L_G \). Then \( L_G^{+}/2 v_2 = 0 \). So
\[ v^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} v = (v_1 + v_2)^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} (v_1 + v_2) \]
\[ = v_1^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} v_1 + 2v_1^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} v_2 + v_2^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} v_2 \]
\[ = v_1^T \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} v_1 \]
\[ \leq v_1^T v_1 = v^T I_{\text{im } L_G} v. \]
Here, the second equality is by the distributive law and the inequality is by (3). Since this holds for every vector \( v \), Fact 4 in the Notes on Symmetric Matrices again implies
\[ \frac{L_G^{+}/2 \cdot X_e \cdot L_G^{+}/2}{r_e} \preceq I_{\text{im } L_G}. \]
Since \( \text{im } X_e \subseteq \text{im } L_G \), Claim 16 in the Notes on Symmetric Matrices shows this is equivalent to
\[ \frac{n-1}{r_e \cdot k} X_e \preceq \frac{n-1}{k} L_G. \]
Since (2) shows that \( E[Z_i] = L_G/k \), this completes the proof of the claim. \( \square \)