Natueal Algorithm **September 17, 2016**

Linear System - Electrical Circuit

Notes: Chi-Ning Chou

In which we discuss the connection between linear system and electrical circuit.

Electrical circuit is an important and interesting physical metaphor of linear system problem. It has a strong and intuitive connection to the graph Laplacian and turned out to be a nice tool for preconditioning. In this post, we study the basic formulation of electrical circuit and see how to use it for sparsification in solving linear system.

1 Preliminary

Definition 1 (incident matrix) *Given a graph $G = (V, E)$, define its incident matrix B with an orientation as follow:*

- *Each row of B stands for an edge $e \in E$ and each column stands for a vertex $v \in V$.*
- *For entry $B_{e,v}$, if v is the starting point of e, then assign 1. If v is the ending point of e, then assign -1. Otherwise, assign 0.*

The incident matrix B has a direct connection with the graph Laplacian of G.

Lemma 2 *Given a graph G with incident matrix B and graph Laplacian L, then $B^T B = L$.*

1.1 Connection between voltage and current

Now, consider an electrical circuit with the same configuration as G with unit resistance put on each edge. Let $\mathbf{c}_{ext}$ be the external current input at each vertex. By two fundamental physics laws, we have the following equations.

Lemma 3 *By Kirchoff's law and Ohm's law, we have*

$$(\textit{Kirchoff's law})\ B^T\mathbf{i} = \mathbf{c}_{ext} \tag{1}$$

$$(\textit{Ohm's law})\ B\mathbf{v} = \mathbf{i} \tag{2}$$

Combine them, we have

$$B^T B\mathbf{v} = \mathbf{c}_{ext} \tag{3}$$

PROOF:

Kirchoff's law refers to the conservation of current at each vertex, *i.e.*, $(B^T\mathbf{i})_v = \sum_{e\in E} B_{v,e}\mathbf{i}_e =$ net current go through vertex $v = (\mathbf{c}_{ext})_v$.

Ohm's law says that the current goes through a wire is proportional to the voltage difference of the two end points with resistance as the multiplicative factor, *i.e.*, $(\mathbf{i})_{(u,v)} = \mathbf{v}_u - \mathbf{v}_v = (B\mathbf{v})_{(u,v)}$. □

Intuition (the connection of voltage and current in electrical circuit)

Intuitively, (3) provides a connection between the voltage and current in a graph electrical circuit. That is, given the voltage in the graph we can compute the external current via multiply by Laplacian matrix. Reversely, given the external current, one can solve the Laplacian linear system to compute the voltage at each vertex, *i.e.*, $\mathbf{v} = L^{+}\mathbf{c}_{ext}$.

Remark 4 *Note that since the Laplacian matrix is singular with uniform vector spanning the null space, the voltage computed by solving linear system can be shifted with a constant factor.*

Remark 5 *When considering weighted graph with matrix W. For edge $e \in E$, assign its resistance $\mathbf{r}_e = 1/W(e,e)$, then the weighted Laplacian becomes $L = B^T W B$ since the Ohm's law in* (2) *becomes*

$$(\textit{weighted Ohm's law})\ B\mathbf{v} = W^{-1}\mathbf{i} \tag{4}$$

1.2 Effective resistance

When we think of an electrical circuit as a gadget, or a digital component, where we inject certain amount of current into the a vertex and allow the same amount of current flowing out from another vertex, what we would care is that what's the *effective resistance* of this gadget? Equivalently, what's the voltage difference of these two end vertices?

From (3), if we inject unit current into vertex i and let it flow out from vertex j, we can compute the voltage with the pseudoinverse of Laplacian matrix:

$$\mathbf{v} = L^{+}(\mathbf{e}_i - \mathbf{e}_j) \tag{5}$$

, where $\mathbf{e}_i$ is the unit vector of the ith entry and note that the voltage $\mathbf{v}$ might be shifted with a constant.

As a result, we know that the voltage difference among vertex i and j is $\mathbf{v}_i - \mathbf{v}_j = (\mathbf{e}_i - \mathbf{e}_j)^T\mathbf{v}$. Thus, by Ohm's law, for any edge $e = (i,j)$, the effective resistance of e is simply

$$(\text{effective resistance})\ R_{eff}(e) = \frac{\mathbf{v}_i - \mathbf{v}_j}{1} = (\mathbf{e}_i - \mathbf{e}_j)^T L^{+}(\mathbf{e}_i - \mathbf{e}_j) \tag{6}$$

More generally, we can compute the voltage difference among any two vertices, say $f = (i', j')$ when having unit current goes into i and goes out from j.

$$R_{eff}(f,e) = (\mathbf{e}_{i'} - \mathbf{e}_{j'})^T L^{+}(\mathbf{e}_i - \mathbf{e}_j) \tag{7}$$

As we can take f and e arbitrarily from E, we can naturally define a matrix Π where $\Pi_{f,e} := R_{eff}(f,e)$ for any $f, e \in E$. Once we think of $(\mathbf{e}_i - \mathbf{e}_j)$ as the e-th row of B, it turns out that we can write Π as follow:

$$\Pi = BL^{+}B^T \tag{8}$$

Here, we summarize several useful properties of Π matrix.

Proposition 6 *1. Π is symmetric.*

2. Π is a orthogonal projection matrix to the column space of B.

3. The eigenvalues of Π are either 0 or 1. Moreover, when G is connected, the multiplicities of 1is exactly $n-1$.

PROOF:

1. As L^+ is symmetric, $\Pi^T = B(L^+)^T B^T = \Pi$.

2. Note that $\Pi^2 = BL^+B^TBL^TB^T = BL^+LL^+B^T = BL^+B^T = \Pi$. Moreover, Π is actually the hat matrix (influence matrix) of B.

3. This is a direct consequence of Π being an orthogonal projection matrix. When G is connected, we know that the rank of B and L are both $n-1$.

□

Finally, there's an important connection between effective resistance and spanning tree on the graph.

Theorem 7 (effective resistance and spanning tree) *Let T be a spanning tree chosen uniformly from all spanning tree of graph G. Then, the probability of an edge e to be chosen in T is*

$$\mathbb{P}[e \in T] = R_{eff}(e) \tag{9}$$

PROOF: There are two ways to prove the theorem. One is algebraic approach and the other is algorithmic.

- (**Algebraic**)
- (**Algorithmic**)
  First, we need the following two lemmas about the random walk on a graph. Proofs of the lemmas are left in the appendix A.1 and A.2.

  Lemma 8 *When enforcing the voltage difference among s and t is 1and shifting the voltage to $\mathbf{v}(s) = 1$ and $\mathbf{v}(t) = 0$, we have $\forall u \in V$,*

  $$\mathbb{P}[\text{random walk starts from } u \text{ reaches } s \text{ before } t] = \mathbf{v}(u) \tag{10}$$

  Lemma 9 *The probability that a random walk starts at s visits t via the edge $e = (s,t)$ is $R_{eff}(e)$.*

  Now, let's randomly pick a spanning tree by the following algorithm: From Lemma 9, we know that the probability of $e = (s,t)$ to be contained in the tree output by the algorithm is exactly $R_{eff}(e)$.

Algorithm 1 Using Random Walk to Uniformly Sample Spanning Tree

1: Perform random walk starting at s.
2: When going out from a vertex v, if v and the visiting edge haven't be chosen, pick this edge.
3: Walk until every vertices are chosen.
4: Output spanning tree.

□

Intuition (effective resistance)

The effective resistance among edge (s,t) has the following intuitions:

- The probability that a random walk starts at s visits t via the edge (s,t).
- The probability that uniformly picking a spanning tree that contains edge (s,t).

1.3 Electrical flows

Max-Flow is one of the most important fundamental problems in computer science in which we consider what's the maximum amount of flow going from source vertex s to terminal t under given capacity constraints. Now, let's consider a variation of this flow problem: Given an unit flow from s which leaving from t, which flow $\mathbf{f}$ can minimize the **flow energy** defined as follow?

$$\text{(flow energy) } E(\mathbf{f}) := \sum_{e\in E} \mathbf{f}_e^2 \tag{11}$$

Not so surprising, the answer is the **electrical flow**! Here, the electrical flow is directly yielded by injecting an unit current in to vertex s and taking the current out from vertex t in graph G where the edges have unit resistance. From the previous subsection, we know that the electrical flow among any two vertices, say i and j, is $(\mathbf{e}_i - \mathbf{e}_j)^T L^+(\mathbf{e}_s - \mathbf{e}_t)$ and thus we can write down the electrical flow $\mathbf{f}^*$ as

$$\text{(electrical flow) } \mathbf{f}^* = BL^+(\mathbf{e}_s - \mathbf{e}_t) \tag{12}$$

The following theorem tells us that the electrical minimize the flow energy defined in (11).

Theorem 10 (electrical flow minimize the flow energy) *Given graph G with unit resistance on each edge, the electrical flow defined in* (12) *minimize the flow energy defined in* (11) *among all unit flows from s to t.*

Proof: First, compute the flow energy of $\mathbf{f}^*$:

$$E(\mathbf{f}^*) = (\mathbf{f}^*)^T\mathbf{f}^* = (\mathbf{e}_s - \mathbf{e}_t)^T(L^+)^T B^T BL^+(\mathbf{e}_s - \mathbf{e}_t) = (\mathbf{e}_s - \mathbf{e}_t)^T L^+(\mathbf{e}_s - \mathbf{e}_t) \tag{13}$$

Now, let $\mathbf{f}$ be arbitrary flow with unit flow injected in s and taken out from t, by Kirchoff's law in (1), we have $B^T\mathbf{f} = (\mathbf{e}_s - \mathbf{e}_t)$ and the energy of $\mathbf{f}$ is $E(\mathbf{f}) = \|\mathbf{f}\|^2$. As Π is an orthogonal projection matrix, we have the following inequality:

$$\|\mathbf{f}\|^2 \geq \|\Pi\mathbf{f}\|^2 = \mathbf{f}^T\Pi^T\Pi\mathbf{f} = \mathbf{f}^T\Pi\mathbf{f} \tag{14}$$
$$= \mathbf{f}^T BL^+B^T\mathbf{f} \tag{15}$$
$$(\because (1)) = (\mathbf{e}_s - \mathbf{e}_t)^T L^+(\mathbf{e}_s - \mathbf{e}_t) \tag{16}$$
$$(\because (13)) = E(\mathbf{f}^*) \tag{17}$$

□

2 Graph Sparsification

The idea of sparsification is simple and straightforward: *reducing the number of edges while in the meantime preserving certain graphical properties.* For instance, given a graph G, finding a sparse **weighted** graph H such that for any cut $(S,\bar{S})$ in G, the weighted number of edges across $S,\bar{S}$ in H is within $1\pm\epsilon$ multiplicative factor in the number of edges in G. Moreover, two quantities here are important for measuring the quality of the sparsification:

- The number of edges in H. Normally, the goal is to find H with $\tilde{O}(\frac{n}{poly(\epsilon)})$ number of edges.
- The time for finding H. Normally, the goal is to find H in $\tilde{O}(m)$ time.

Here, what we are interested is the **spectral sparsification** of graph G which preserve the Laplacian norm in the following sense.

Definition 11 (ϵ-spectral sparsifier) *Given an undirected graph G, a weighted graph H is said to be an ϵ-spectral sparsifier if* $\forall\mathbf{x}\in\mathbb{R}^n$

$$\frac{1}{1+\epsilon} \leq \frac{\mathbf{x}^T L_H\mathbf{x}}{\mathbf{x}^T L_G\mathbf{x}} \leq 1+\epsilon \tag{18}$$

One can see that when H is a spectral sparsifier of G, the solution of $L_H\mathbf{x}=\mathbf{b}$ will be close to the solution of $L_G\mathbf{x}=\mathbf{b}$ with some ϵ factor. As long as $L_H\mathbf{x}=\mathbf{b}$ can be solved with a fast algorithm, it is a nice tradeoff.

In the following section, we are going to see how to use the effective resistance in electrical circuit to construct a good spectral sparsifier efficiently.

2.1 Spectral sparsification using effective resistance

The high-level motivation of using effective resistance to construct spectral sparsification is based on Theorem 7. Somehow, a spanning tree in G turns out to be a good spectral sparsifier. In later posts, we will have a look at the historical development of this brilliant idea.

Here, we assume the effective resistance R_{eff} in the given graph G is known. In later extension, these quantity needed to be estimated as well. The following is the idea of constructing spectral sparsifier with effective resistance.

1. Sample T edges $Y_1,\dots,Y_T$ with probability $\mathbb{P}[Y_i=e]\sim R_{eff}(e)\ \forall i\in[T]$.

2. Using $Y_1, \dots, Y_T$ to construct an estimator $\tilde{\Pi}$ of Π and show that they are similar.
3. Show that the maximum spectral ratio can be upper bounded by the difference of Π and $\tilde{\Pi}$

Next, let's see how to construct and analyze the spectral sparsifier using effective resistance.

2.1.1 Sampling edges using effective resistance

To sample $Y_1, \dots, Y_T$ with probability proportional to the effective resistances, we need to first know the sum of the effective resistances so that we can normalize them. From 3. of Proposition 6, we know that the sum of eigenvalues are $n-1$. As long as $\sum_e R_{eff}(e) = tr(\Pi) = \sum_i \lambda_i = n-1$, we can define the sampling probability as follow:

$$p_e := \frac{R_{eff}(e)}{n-1} \tag{19}$$

Furthermore, let the resulting subgraph of G to be H, and assign each edge with weight $w(e) = \frac{1}{\sqrt{p_e}}$, we can show that L_H is an unbiased estimator of L_G. That is,

$$\begin{aligned}
\mathbb{E}[L_H] &= \frac{1}{T}\sum_{i\in[T]}\mathbb{E}[\frac{\mathbf{b}_{Y_i}}{\sqrt{p_{Y_i}}}\frac{\mathbf{b}_{Y_i}^T}{\sqrt{p_{Y_i}}}] && (20)\\
&= \frac{1}{T}\sum_{i\in[T]}\sum_{e} p_e\frac{\mathbf{b}_e}{\sqrt{p_e}}\frac{\mathbf{b}_e^T}{\sqrt{p_e}} && (21)\\
&= \frac{1}{T}\sum_{i\in[T]}\sum_{e}\mathbf{b}_e\mathbf{b}_e^T = L_G && (22)
\end{aligned}$$

2.1.2 Constructing estimator of Π

Recall that $\Pi := BL_G^+B^T$, we construct an estimate of Π with $Y_1, \dots, Y_T$ as follow:

$$\tilde{\Pi} := \frac{1}{T}\sum_{i\in[T]}\mathbf{v}_{Y_i}\mathbf{v}_{Y_i}^T \tag{23}$$

, where $\mathbf{v}_e = \frac{\Pi_e}{\sqrt{p_e}}$ and Π_e is the column of Π corresponding to edge e. Similarly as the Laplacian induced by $Y_1, \dots, Y_T$, $\tilde{\Pi}$ is also an unbiased estimator of Π. Next, we are going to apply the following lemma to show the concentration of Π. Note that the matrix norm we use here is the *operator norm, i.e.,* $\|M\| := \sup_{\mathbf{x}\neq 0}\frac{\mathbf{x}^TM\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$.

Lemma 12 *Let $\epsilon > 0$ and $M_1, \dots, M_T \in \mathbb{R}^{d\times d}$ be T i.i.d. copies of random matrices with expectation $\mathbb{E}[M_i] = I_d$, the identity matrix in d-dimensional space. Let $\rho := \sup_M \|M\|$, we have*

$$\mathbb{P}[\|\frac{1}{T}\sum_{i\in[T]}M_i - I_d\| > \epsilon] \leq 2d\cdot\exp\Big(-\frac{T\epsilon^2}{\rho}\Big) \tag{24}$$

Remark 13 *Lemma 12 holds for M whose expectation is unitarily equivalent to I_d.*

As for any $\mathbf{v}_e$, $\|\mathbf{v}_e\|^2 \leq \frac{R_{eff}(e)}{p_e} = n-1$, we have $\rho = n-1$. Apply Lemma 12 to $\tilde{\Pi}$, we have

$$\mathbb{P}[\|\tilde{\Pi}-\Pi\| > \epsilon] \leq 2d\cdot\exp\Big(-\frac{T\epsilon}{2(n-1)}\Big) \tag{25}$$

2.1.3 Upper bound the spectral ratio via $\|\tilde{\Pi}-\Pi\|$

Finally, observe the relation of $\tilde{\Pi}$ and Π as follow:

$$\tilde{\Pi} = \frac{1}{T}\sum_{i\in[T]} \mathbf{v}_{Y_i}\mathbf{v}_{Y_i}^T \tag{26}$$

$$= \frac{1}{T}\sum_{i\in[T]} \frac{\Pi_{Y_i}\Pi_{Y_i}^T}{p_e} \tag{27}$$

$$(\because \Pi_e = BL_G^+\mathbf{b}_e) = \frac{1}{T}\sum_{i\in[T]} \frac{1}{p_e}(BL_G^+\mathbf{b}_e\mathbf{b}_e^T L_G^+ B^T) \tag{28}$$

$$(\because L_H = \frac{1}{T}\sum_{i\in[T]} \frac{\mathbf{b}_{Y_i}}{\sqrt{p_{Y_i}}}\frac{\mathbf{b}_{Y_i}^T}{\sqrt{p_{Y_i}}}) = BL_G^+ L_H L_G^+ B^T \tag{29}$$

That is, $\tilde{\Pi}-\Pi = BL_G^+(L_H - L_G)L_G^+B^T$, and we have

$$\|\tilde{\Pi}-\Pi\| = \sup_{\mathbf{x}\neq 0}\Big|\frac{\mathbf{x}^T BL_G^+(L_H-L_G)L_G^+B^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\Big| \tag{30}$$

$$(\text{subsitute } \mathbf{x} \text{ with } B\mathbf{z}) \geq \sup_{\mathbf{z}\neq 0, \langle z,\mathbf{1}\rangle = 0}\Big|\frac{\mathbf{z}^T B^T BL_G^+(L_H-L_G)L_G^+B^T B\mathbf{z}}{\mathbf{z}^T B^T B\mathbf{z}}\Big| \tag{31}$$

$$(\because L_G = B^T B) = \sup_{\mathbf{z}\neq 0, \langle z,\mathbf{1}\rangle = 0}\Big|\frac{\mathbf{z}^T(L_G-L_H)\mathbf{z}}{\mathbf{z}L_G\mathbf{z}}\Big| \tag{32}$$

Combine (29) and (32), we have

$$\mathbb{P}[\sup_{\mathbf{z}\neq 0,\langle\mathbf{z},\mathbf{1}\rangle} > \epsilon] \leq \mathbb{P}[\|\tilde{\Pi}-\Pi\| > \epsilon] \leq 2d\cdot\exp(-\frac{T\epsilon^2}{2(n-1)}) \tag{33}$$

Take $T = \tilde{O}(\frac{n}{\epsilon^2})$, we have an ϵ-spectral sparsifier for G.

A Proofs

A.1 Proof of Lemma 8

Proof: Denote the probability as $p_{s,t}(u)$ we have

$$p_{s,t}(u) = \sum_{v\sim u}\mathbb{P}[u\to v]p_{s,t}(v) = \sum_{v\sim u}\frac{1}{deg(u)}p(v) \tag{34}$$

with boundary condition $p_{s,t}(s) = 1$ and $p_{s,t}(t) = 0$. It turns out that (34) is equivalent to the Kirchoff's law in (1) and the boundary conditions of $\mathbf{v}$ and $p_{s,t}$ are the same. Thus, $\mathbf{v} = p_{s,t}$ and the lemma holds. □

A.2 Proof of Lemma 9

PROOF: Let x be the desired probability, we have

$$x = \mathbb{P}[s \to t] + \sum_{v \sim s, v \neq t} \mathbb{P}[s \to v] \cdot p_{s,t}(v) \cdot x \tag{35}$$

$$= \frac{1}{deg(s)} + \sum_{v \sim s, v \neq t} \frac{x}{deg(s)} \cdot p_{s,t}(v) \tag{36}$$

Solving x we have

$$x = \frac{1/deg(s)}{1 - \sum_{v \sim s, v \neq t} p_{s,t}(v)/deg(s)} = \frac{1}{deg(s) - \sum_{v \sim s, v \neq t} p_{s,t}(v)} \tag{37}$$

$$= \frac{1}{\sum_{v \sim s, v \neq t} 1 - p_{s,t}(v)} \tag{38}$$

$$(\because p_{s,t}(s) = 1) = \frac{1}{\sum_{v \sim s, v \neq t} p_{s,t}(s) - p_{s,t}(v)} \tag{39}$$

$$(\because \textit{Lemma } 8) = \frac{1}{\sum_{v \sim s} \mathbf{v}(s) - \mathbf{v}(t)} \tag{40}$$

$$(\because \text{Kirchoff's law}) = \frac{1}{\text{current flows out } s} \tag{41}$$

$$(\because \text{Ohm's law}) = R_{eff}(e) \tag{42}$$

□

References