15-451: Algorithms December 3, 2020

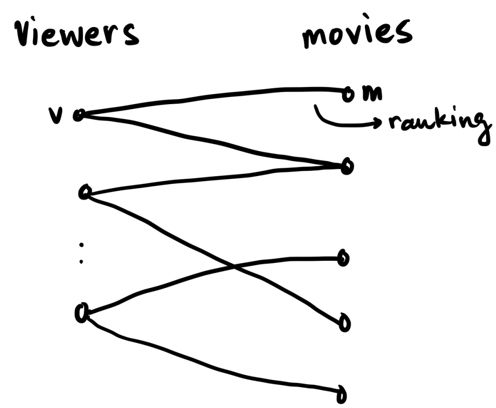
Lecture Notes: Resistive Model of a Graph and Random Walks

Lecturer: Gary Miller *Scribes:*

[1]

0 Motivation

Consider a recommendation system that recommends viewers to movie titles. We can represent this system as a bipartite graph $G = (L, R, E)$ with a ranking function r where L is the set of viewers, R is the set of movies, E is the set of edges, and $r : E \to \mathbb{N}$ gives for $(v, m) \in E$ viewer v's ranking of movie m.



Using this model, we want to answer the following question: Should we recommend movie m to viewer v? In other words, how can we assign some sort of score score(v, m) to a recommendation of m to v. (**Note:** (v, m) need not be an edge in G - we want to be able to generate these recommendation scores regardless of whether a v has a ranking for m). Let's explore a couple of ideas.

[1]Originally 15-859N notes by Tanvi Bajpai

Idea 1

$$\text{score}(v,m) = \frac{1}{\text{dist}_G(v,m)}$$

where we define $\text{dist}_G(v,m)$ as follows: Assign a weight to each edge (i,j) in the graph given by

$$w_{ij} = \frac{1}{r(i,j)}$$

and the weight of a path P is given by

$$w(P) = \sum_{e \in P} w_e$$

Now, let

$$\text{dist}_G(v,m) = \min_{vPm} W(P)$$

Idea 2

$$\text{score}(v,m) = \max_{vPm} w(P)$$

where we define

$$w(P) = \min_{e \in P} r(e)$$

However, these two ideas don't help us since we never take into account multiple paths between v and m to add to the score. Intuitively, the more paths there are between a viewer and a movie, the higher the $\text{score}(v,m)$ should be. We can amend this by considering the following idea:

Idea 3

$$\text{score}(v,m) = \text{ max flow from } v \text{ to } m$$

This is *still* not good enough, since max flow won't reward shorter paths between viewers and movies. This motivates two other ideas:

Idea 4

View the edges as conductors, and let

$$\text{score}(v,m) = \text{hit}(v,m) + \text{hit}(m,v)$$

where $\text{hit}(v,m)$ denotes the expected length of a random walk from v to m

Idea 5

Consider a random walk from v to m

$$\text{score}(v, m) = \text{ “the effective conductance” between } v \text{ and } m$$

We will show that Ideas 4 and 5 are equal up to scaling. We will also explore whether effective conductance and “commute time” provide a better scoring system.

The rest of the lecture will proceed as follows:

- Provide formal definitions
- Develop a Basic Theory
- Give efficient algorithms
- Find applications

1 Resistance Theory

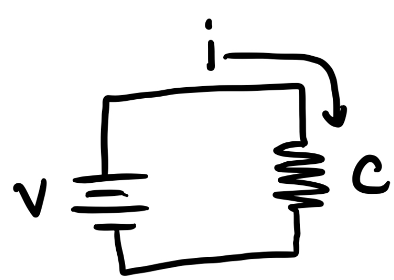
1.1 Some preliminaries

We'll first use some basic laws of physics adn shit

Ohm's Law

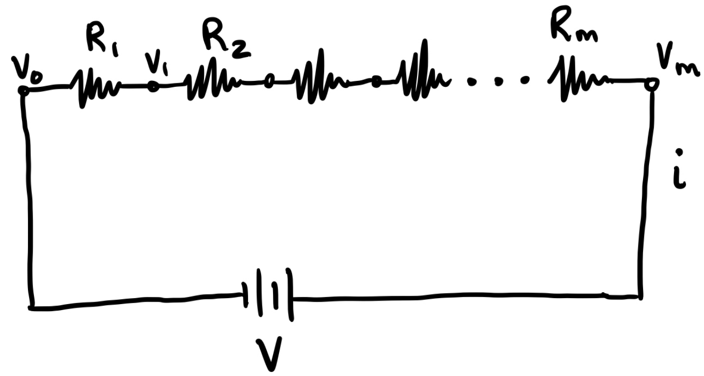
Let $C =$ conductance, $R =$ resistance, $V =$ voltage, and $i =$ current, then

$$C = \frac{1}{R} \qquad i = CV = \frac{V}{R}$$



Resistors in Series

Series of resistors will act like a single resistor



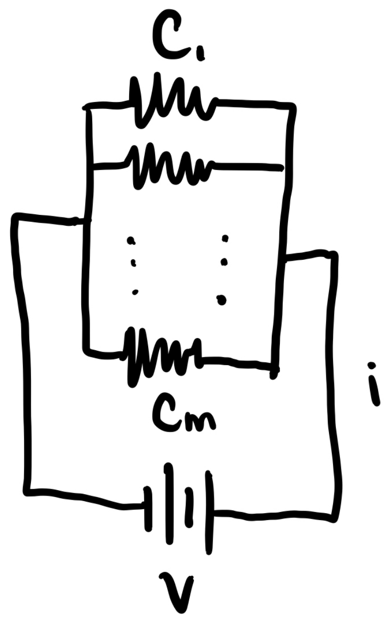
$$R = R_1 + ... + R_n$$

$$C = \frac{1}{(1/C_1 + 1/C_2 + ... + 1/C_n)}$$

$$i = \frac{V}{R}$$

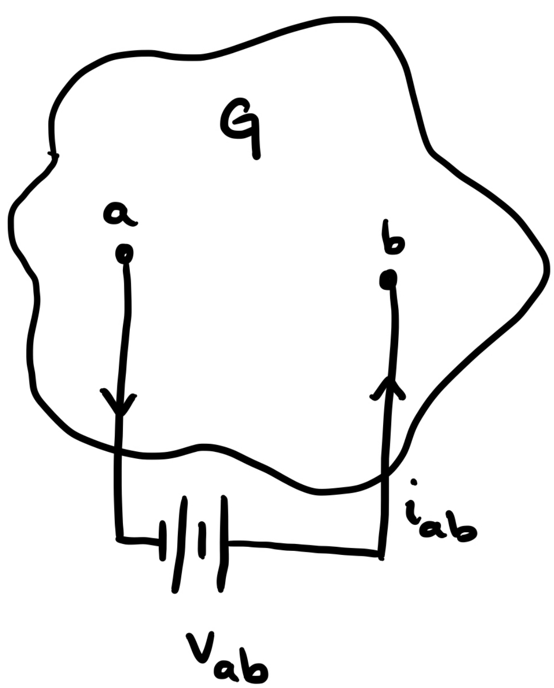
Conductors in Parallel

Conductors in parallel will act like a single conductor



$$C = C_1 + ... + C_m$$

1.2 Effective Resistance



Definition 1. *Let G be a network of resistors, where the conductance on each individual edge is given by the edge weight. The* ***Effective Resistance (Conductance)*** *between two vertices a and b of in G is the amount of electrical resistance (conductance) between them:*

$$R_{ab} = \frac{V_{ab}}{i_{ab}} \qquad C_{ab} = \frac{1}{R_{ab}}$$

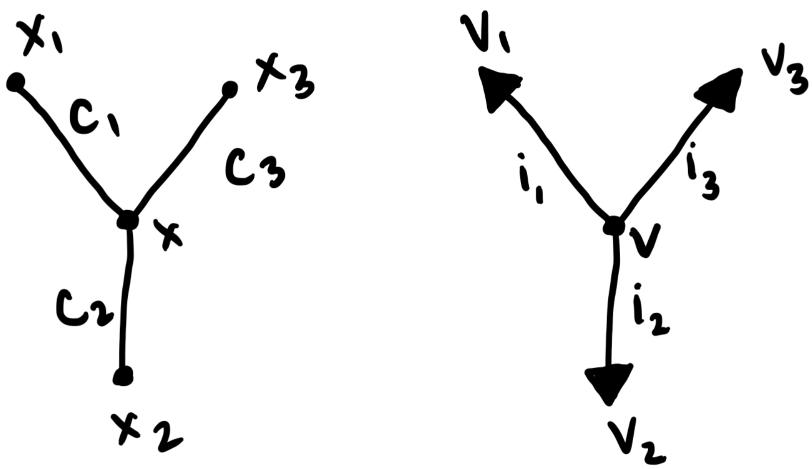
Now, to compute the effective resistance from a to b, we will make use of **Kirchoff's Law**

Kirchoff's (1st) Law

The current flowing into a node must be equal to the current flowing out of it.

(This will be the case for all nodes other than a and b)

Now, consider the following example (the left is the weighted graph, while the right represents the voltage and current flow)



Using Ohm's Law, we have that

$$\begin{aligned} i_1 &= C_1(V - V_1) \\ i_2 &= C_2(V - V_2) \\ i_3 &= C_3(V - V_3) \end{aligned}$$

The residual current is given by $i_1 + i_2 + i_3$. Kirchoff's Law tells us that this quantity must equal 0, hence

$$C_1(V - V_1) + C_2(V - V_2) + C_3(V - V_3) = 0 \implies (C_1 + C_2 + C_3)V = C_1V_1 + C_2V_2 + C_3V_3$$

Setting $C = C_1 + C_2 + C_3$ gives us

$$CV = C_1V_1 + C_2V_2 + C_3V_3 \implies V = \frac{C_1}{C}V_1 + \frac{C_2}{C}V_2 + \frac{C_3}{C}V_3$$

Observe that we've written V as a convex combination of V_1, V_2, V_3. The residual current is equal to $CV - C_1V_1 - C_2V_2 - C_3V_3$

Now, we'll consider the general case for some arbitrary network G, where $G = (V, E, C)$ where $C : E \to \mathbb{R}^+$ and $V = \{1, \ldots, n\}$. Also, define

$$d(a) = \sum_{(a,b)\in E} C(a, b)$$

Define a corresponding adjacency matrix A whose entries are given by

$$A_{ab} = \begin{cases} C(a, b) & (a, b) \in E \\ 0 & o.w. \end{cases}$$

The Laplacian matrix of G (denoted by $L(G)$ or simply L) is given by

$$L_{ab} = \begin{cases} d(a) & a = b \\ -C_{ab} & (a, b) \in E \\ 0 & o.w. \end{cases}$$

Notice that

$$L = D - A \qquad \text{where } D = \begin{bmatrix} d(1) & & 0 \\ & \ddots & \\ 0 & & d(n) \end{bmatrix}$$

Let $\mathbf{v}$ be a vector representing the voltage setting of each node; This would mean that $(L\mathbf{v})_i$ calculates the residual current at node i. Now, suppose we are interested in the inverse, i.e. we inject current into each node and observe the voltage. The net current injected should be zero (in order to abide by Kirchoff's law). All we'd need to do is solve for $L\mathbf{v} = \mathbf{i}$ (where $\mathbf{i}$ represents the vector of currents)

Now, let us try computing the value of the effective resistance between nodes 1 and n (i.e. R_{1n}. We can approach this computation in two different ways

Method 1: Solve

$$(*) \quad L \begin{bmatrix} 1 \\ V_2 \\ \vdots \\ V_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} i \\ 0 \\ \vdots \\ 0 \\ -i \end{bmatrix} \qquad i = \frac{V}{R}, \quad V = 1, \quad R = \frac{1}{i}$$

and return $\frac{1}{i}$

$(*)$ This is called a boundary valued problem. In our case, V_1 and V_n are the boundary; $(V_1, ..., V_n)$ is called harmonic, since all of the interior V_a can be written as a convex combination of its neighbors.

Maximum Principle: If f is harmonic, then its min and max are on the boundary.

Proof. If V is interior, then there must exists neighbors V_a and V_b such that $V_a \leq V \leq V_b$. $\square$

Uniqueness Principle: If f and g are harmonic with the same boundary values, then $f = g$

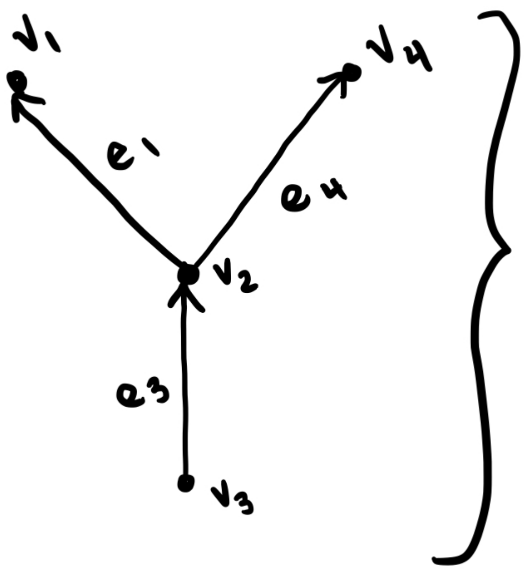
Proof. $f - g$ is harmonic, with zero's at both boundaries, hence $f - g \equiv 0$, therefore $f = g$. $\square$

Method 2: Solve

$$L\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}$$

and return $R_{1n} = V_1 - V_n$ (How do we know that an assignment of $\mathbf{v}$ exists?)

Another way to view the Laplacian: Boundary Operator (vertex-edge matrix): $B^{n \times m}$ and pick a direction to orient each edge. Consider the following example



$$B \equiv \begin{array}{c|ccc} & e_1 & e_3 & e_4 \\ \hline v_1 & 1 & 0 & 0 \\ v_2 & -1 & 1 & -1 \\ v_3 & 0 & -1 & 0 \\ v_4 & 0 & 0 & 1 \end{array}$$

Let $C_1, ..., C_m$ denote the conductances of $e_1, ..., e_m$

$$\mathbf{C} = \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_m \end{bmatrix}$$

Note that if f is a flow then Bf will equal, for each vertex, the surplus flow at that vertex. Thus $-Bf$ will be the needed injected flow at each vertex.

Observe that

- $B^T\mathbf{v} \equiv$ the voltage drop across each edge
- $\mathbf{C}B^T\mathbf{v} \equiv$ **minus** the current flow along each edge
- $\mathbf{C} - B^T\mathbf{v} \equiv$ the current flow along each edge by Ohm's Law.
- $(-B)\mathbf{C}(-B^T\mathbf{v}) \equiv$ the need injected current at each vertex

This gives us that $L = B\mathbf{C}B^T$

If G is a connected graph, we're interested in answering the following questions:

- What is rank(L)?
- What is ker(L)?

Consider

$$x^T Lx = x^T B\mathbf{C}B^T x = (B^T x)^T \mathbf{C}B^T x = \sum_{(a,b)\in E} \mathbf{C}_{ab}(x_a - x_b)^2$$

thus

$$x^T L = 0 \iff \forall (a,b) \in E \;\; (x_a - x_b)^2 = 0 = (x_a - x_b)$$

Therefore if G is connected $\implies \forall a, b \;\; x_i = x_j$

Hence, the kernel of L is $\langle \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle$ while the rank is $n - 1$

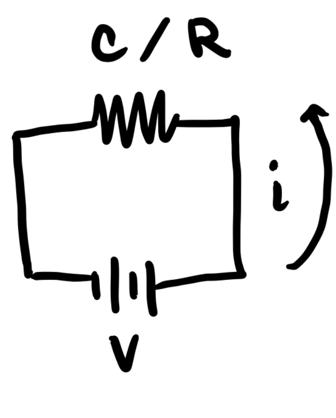
Claim 1. *$Lx = 0$ iff $x^T Lx = 0$*

Proof. The forward direction is clear, the backwards direction is

$$x^T Lx = x^T BCB^T x = (\mathbf{C}^{1/2}B^T x)^T(\mathbf{C}^{1/2}B^T x) = 0$$

which means that $\mathbf{C}^{1/2}B^T x = 0$ which means that $B\mathbf{C}^{1/2}\mathbf{C}^{1/2}B^T x = 0$, hence $Lx = 0$ □

1.3 Current and Energy/Power Dissipation



Newton said that

$$\begin{aligned}\text{Energy} &\equiv \text{Force} \cdot \text{Speed} \\ &\equiv \text{Volt} \cdot \text{Current} \\ &\equiv Vi \\ &\equiv CV^2 \equiv i^2 R\end{aligned}$$

In our network, this means that Energy $= \frac{1}{2} \sum_{x,y} |i_{xy}| \cdot |\mathbf{v}_x - \mathbf{v}_y|$
$v^T L v = v^T B \mathbf{C} B^T v = (B^T v)^T \mathbf{C} (B^T v) = \sum_{\text{oriented}(x,y) \in E} \mathbf{C}_{xy} (V_x - V_y)^2 = \text{Energy}$

1.4 Flows

We define two types of flows in our network:

Definition 2. *A **flow** is a function $f : E \to \mathbb{R}$ over oriented edges*

Definition 3. ***potential flow*** $= \{\boldsymbol{C} B^T \boldsymbol{v} \mid \boldsymbol{v} \in \mathbb{R}^n\} \equiv P_G$

Definition 4. ***circulation flows*** $= \{f \in \mathbb{R}^m \mid Bf = 0\} \equiv C_G$

Assume G is connected and we are given its spanning tree T.

Claim 2. *C_G is a subspace, and the* $\dim(C_G) = m - n + 1$

Note that:

- Proving that this is a subspace is easy.
- $E \setminus T$ denotes the non-tree edges of G, hence $|E \setminus T| = m - n + 1$
- Any flow on $E \setminus T$ can be extended to C_G on G (to be shown in homework)
- $f, g \in C_G$ and $f \setminus T = g \setminus T$ then $f = g$

Claim 3. *The dimension of the subspace P_G is $n - 1$*

Claim 4. *$f_C \in C_G$ and $g_P \in P_G$ then $f_C^T R g_P = 0$ where* $R = \begin{bmatrix} R_1 & & 0 \\ & \ddots & \\ 0 & & R_m \end{bmatrix}$

Proof. $\exists v$ such that $g_P = C B^T \mathbf{v}$. Now we have

$$f_C^T R g_P = f_C^T R C B^T \mathbf{v} = f_C^T B^T \mathbf{v} = (B f_C)^T \mathbf{v} = 0^T \mathbf{v} = 0$$

(observe that we use the fact that $RC = I$) □

Therefore, $C_G; P_G$ spans $\mathbb{R}^m$ (all flows) i.e. $\forall f \in \mathbb{R}^m$, $\exists! f_C, f_P$ such that $f = f_C + f_P$

Definition 5. *$f_a = \sum_{a \neq b} f_{ab}$ where $a, b \in V$*

Definition 6. *f is a **unit flow** from a to b if:*

- *f is a flow*
- *$f_a = 1$ and $f_b = -1$*
- *$f_x = 0$ for $x \neq a, b$*

Thomson's Principle: If the following two conditions hold, then $f^T R f \leq g^T R g$

- f is unit potential flow from a to b
- g is any unit flow from a to b

Proof. We know that $g = f + f_c$ where $f_C \equiv$ circulation

$$g^T R g = (f + f_C)^T R(f + F_C) = f^T R f + 2 f_C R f + f_C^T R f_C = f^T R f + f_C^T R f_C \geq f^T R f$$

□

Thus Thomson Principle gives yet another equivalent definition effective resistance is:

Definition 7. *The effective resistance from a to b (ER_{ab}) can be defined as $f_P^T R f_p$, where f_p denotes the unit potential flow from a to b.*

Rayleigh's Monotonicity Law: If $\overline{R} \geq R$ then $\overline{ER}_{ab} \geq ER_{ab}$

Proof. Let f be the unit potential flow in G_R, and g be the unit potential flow in $G_{\bar{R}}$

$$\begin{aligned} \overline{ER}_a b = g^T \overline{R} g &= \sum_{e \in G} g_e^2 \overline{R}_e \\ &\geq \sum_{e \in G} g_e^2 R_e \\ &\geq \sum_{e \in G} f_e^2 R_e \qquad \text{(by Thomson)} \\ &= f_e^T R f_e = ER_{ab} \end{aligned}$$

□

These two results allow us to show that R_{ab} is a metric. That is:

HW: Show that $R_(ab)$ is a metric space, i.e.,

- $R_{ab} \geq 0$
- $R_{ab} = 0$ iff $a = b$
- $R_{ab} = R_{ba}$
- $R_{ac} \leq R_{ab} + R_{bc}$

2 Random Walks

Let $G = (V, E, w)$ be a (possibly directed) graph where

$$w_a = w(a) \equiv \sum_{(a,b)\in E} w_{ab}$$

$$P_{ab} \equiv \frac{w_{ab}}{w_a}$$

Definition 8. ***Random walk*** *on G: Suppose at a given time we are at $a \in V$, we move to b with probability P_{ab}*

Ex. Let V be all orderings of a deck of 52 cards. P_{ab} will be the probability of going from some order a to an order b in one shuffle.

Fun question: Why do professionals play after 5 shuffles?

We can consider two views of a random walk:

- Particle view (the definition above)
- Wave view: there's a large number of simultaneous independent walkers

$$x^{(i)} \equiv \text{distance at time } i$$

$$x^{(i+1)} = AD^{-1}x^{(i)}$$

Definition 9. ***Access Time****, or* ***Hitting Time*** *H_{ab} is the expected time to visit b starting at a.*

Definition 10. ***Commute Time:*** *$K_{ab} = H_{ab} + H_{ba}$*

Definition 11. ***Cover Time*** *is the expected time to visit all nodes (we take the max over all starting nodes)*

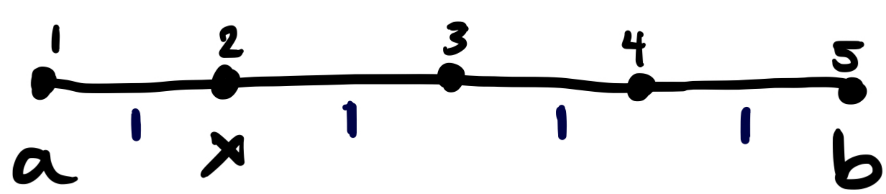
Definition 12. ***Mixing Rate:*** *TODO in a future lecture*

2.1 Random Walks on Symmetric Graphs

Idea: View a random walk as a walk on a network of conductors.

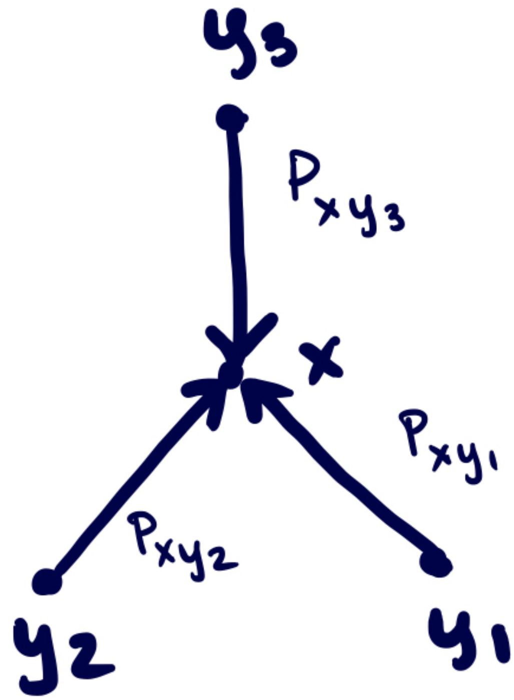
Input: $G = (V, E, C)$ where $C(a, b) = C(b, a)$

Consider a random walk starting at x and ending at b. Let h_x be the probability we visit a before b when starting at x, where $a \neq b$. Consider the following example:



We know that $h_a = 1$ and $h_b = 0$. What if we want to compute h_2? We know that this quantity must be greater than 0.5. After some calculations, we can observe that $h_x = 0.75$.

Example 2:

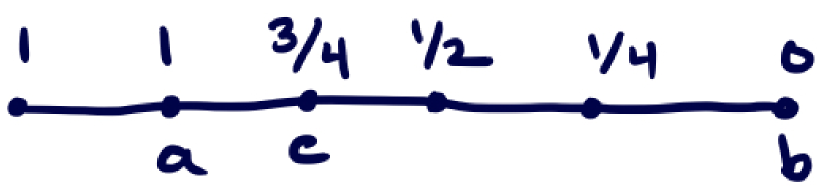


Claim 5. $h_x = \sum_y P_{xy} h_y$

We know that $P_{xy} \geq 0$ and that $\sum_y P_{xy} = 1$. This means that h_x is a convex combination of its neighbors.(Also, h is harmonic with boundary a, b!).

Now, lets construct an identical electrical problem. Consider $V_a = 1$ and $V_b = 0$. $\forall x \neq a, b$, $V_x = \sum_y \frac{C_{xy}}{C_x} V_y$. Observe that $\frac{C_{xy}}{C_x} = P_{xy}$, which means h and V equal.

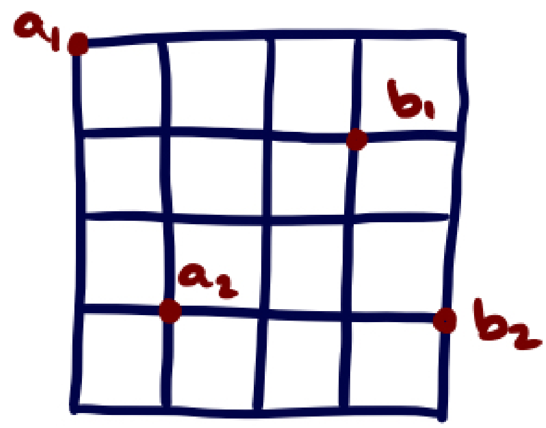
Theorem 1. *Set $V_a = 1$ and $V_b = 0$; Let $x \neq a, b$ "float" and then let V_x be the probability that we visited a before v. The residual current at x will be* 0.



In the above example, h_c equals 0.75, and $a = v_1$ while $b = v_n$. Algebraically we get

$$L \begin{bmatrix} 1 \\ * \\ \vdots \\ * \\ 0 \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \\ * \end{bmatrix}$$

In general, we can have multiple sinks and goals:



We can compute this with one Laplacian solve.

2.2 Interpretation of Current as Random Walk

Consider one unit of potential current flow from a to b, say i. What does i_{xy} correspond to in a random walk from a to b?

Theorem 2. *i_{xy} will be the expected net number of traversals of edge (x, y) in a random walk from a to b.*

Proof. Let U_x be the expected number of visits to x before reaching b starting at a. For HW, show that $\sum_y U_y P_{yx}$ (Note: $\sum_y P_{yx} \neq 1$). Now, recall that $C_x = \sum_y C_{xy}$, and note that

$$C_x P_{xy} = C_x \left(\frac{C_{xy}}{C_x} \right) = C_{xy} = C_{yx} = C_y \left(\frac{C_{yx}}{C_y} \right) = C_y P_{yx}$$

Thus,

$$U_x = \sum_y U_y \frac{C_y P_{yx}}{C_y} = \sum_y U_y \frac{P_{xy} C_x}{C_y}$$

Therefore, $\frac{U_x}{C_x} = \sum_y P_{xy} \left(\frac{U_y}{C_y} \right)$

Here, the voltage $V_x = \frac{U_x}{C_x}$, and its recurrence is $V_x = \sum_y P_{xy} V_y$ (and the residual current at x is 0). This means that V_x is harmonic with boundary conditions: $V_b = 0$ and $V_a = \frac{U_a}{C_a}$ for the "correct" U_a

Define j_{xy} as the current on edge (x, y), and observe that

$$j_{xy} = (V_x - V_y)C_{xy} = \Big(\frac{U_x}{C_x} - \frac{U_y}{C_y}\Big)C_{xy} = U_x\Big(\frac{C_{xy}}{C_y}\Big) = U_x P_{xy} - U_y P_{yx}$$

Here, $U_x P_{xy}$ is the expected number of traversals from x to y (similarly, $U_y P_{yx}$ is the expected number of traversals from y to x). This means that j_{xy} is the expected net number of traversals from x to y.

Now, we wish to show that the net current flow is 1, i.e. $\sum_y j_{ay} = 1$. This value must be 1, since we must have had to leave a once, for good, to get to b (which means we never traveled that edge back).

$\square$

Now, we are interested in computing U_a. Consider a as the first vertex, while b is the nth vertex. Solve

$$L\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ -1 \end{bmatrix}$$

Set

$$\mathbf{v}' = \mathbf{v} - V_n \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

i.e. set $V_n = 0$. Hence $V_1 = \frac{U_1}{C_1} \implies U_1 = V_1 C_1$, which means we have found U_a! Also, $V = ER_{ab}$ and $U_a = C_a \cdot ER_{ab}$

2.3 How to compute hitting time

Recall that H_{xb} denotes the expected time to reach b from x. Let $H_x = H_{xb}$ for some fixed b. Consider the following recurrence:

$$\begin{aligned} H_b &= 0 \\ H_x &= 1 + \sum_y H_y P_{xy} \qquad (x \neq b) \end{aligned}$$

We can think of H_x as a voltage V_x:

$$\begin{aligned} V_b &= 0 \\ V_x &= 1 + \sum_y \Big(\frac{C_{xy}}{C_x}\Big) V_y \end{aligned}$$

$$C_x V_x = C_x + \sum_y C_{xy} V_y$$
$$C_x V_x - \sum_y C_{xy} V_y = C_x$$

Observe that the LHS is the graph Laplacian, while the RHS is the residual current. There are $n-1$ constraints, and by adding constraint for $v_n = b$

$$L\mathbf{v}- = \begin{bmatrix} C_1 \\ \vdots \\ C_{n-1} \\ \delta \end{bmatrix}$$
$$V_n = 0 \qquad C = \sum_i C_i$$

where $\delta = C_n - C$

Now, if we wish to compute the hitting time from a v_x to v_n, solve the above equation for V_x.

To solve for the commute time between a (v_1) and b (v_n), we have two methods:

1. Solve the following two

$$L\mathbf{v}^b = \begin{bmatrix} C_1 \\ \vdots \\ C_n - C \end{bmatrix} \qquad L\mathbf{v}^a = \begin{bmatrix} C_1 - C \\ \vdots \\ C_n \end{bmatrix}$$

   $H_{1n} = V_1^b - V_n^b$ and $H_{n1} = V_n^a - V_1^a$. Set $\mathbf{v} = \mathbf{v}^b - \mathbf{v}^a$ (and so $V_i = V_i^b - V_i^a$), and return

$$K_{1n} = H_{1n} + H_{n1} = V_1 - V_n$$

2. Solve the following

$$L(\mathbf{v}^b - \mathbf{v}^a) = L\mathbf{v}^b - L\mathbf{v}^a$$

   This is equal to

$$\begin{bmatrix} C_1 \\ \vdots \\ C_n - C \end{bmatrix} - \begin{bmatrix} C_1 - C \\ \vdots \\ C_n \end{bmatrix} = \begin{bmatrix} C \\ \vdots \\ -C \end{bmatrix} = C \begin{bmatrix} 1 \\ \vdots \\ -1 \end{bmatrix}$$

   Then, solve for

$$L\mathbf{v} = \begin{bmatrix} 1 \\ \vdots \\ -1 \end{bmatrix}$$

   And return $C(V_1 - V_n)$ where $(V_1 - V_n) = ER_{1n}$

Method 2 motivates the following theorem:

Theorem 3. $K_{ab} = C \cdot ER_{ab} = 2m \cdot ER_{ab}$

For trees, this means that the commute time between two vertices is $2(n-1) \cdot ER_{ab}$