Random Walks and Algorithms

Note Title 4/28/2020

Last time we saw an algebraic view of graphs

$G \longrightarrow$ Adjacency matrix $A \longrightarrow$ Laplacian matrix $L = D - A$.

We also saw that the spectral gap $\lambda_2 - \lambda_1$ (where $0 = \lambda_1(A) \leq \lambda_2(A) \ldots \leq \lambda_n$) indicates how "connected" the graph is.

Today we'll see some other connections:—

① Random Walks on graphs

Graph $G = (V, E)$ maybe undirected unweighted (for now)

We have a token at a vertex of G, each time we pick a random neighbor of it, and move to it.

So X_t = location @ t

then $X_{t+1} \in$ uniformly random one of $N(X_t)$.

Q1 How long to see all vertices of G (Cover time)

Q2 How long to get to specific target vertex v^*. (Hitting time)

Q3: What is the limiting distribution of X_T as $T \to \infty$?

Q4 Why do we care? ☺

Q4 first:

Because G could be a large structure (exponentially sized) and we want to understand some random process on it.

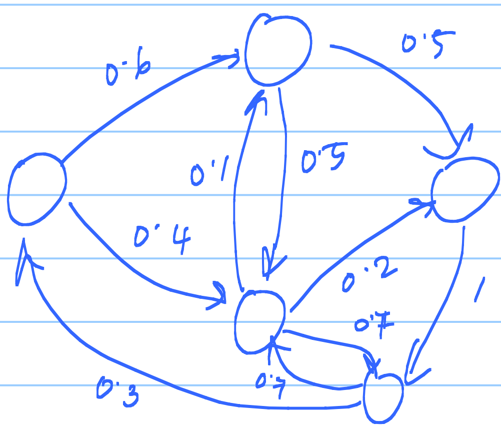
E.g. (Distributed) agents in an unknown environment
How to get out of a maze without a map.

- Target vertex = exit
- Cover time = time to map entire location

Eg: graph is set of all independent sets of G.
Each edge connects two similar independent sets
Time to reach a target = large independent set?

Actually a reverse of Q3 is interesting, because often we want to sample from a given distribution, and want to design a random walk like process that gets us there. (Q5: How fast to reach close to stationary)

So design a (often edge weighted, directed) graph G
Also called a Markov Chain
Such that the limiting distribution of X_T is the distribution we want.



Same kinds of questions on this.

But first: undirected unweighted graphs G and random walks. Let d_u = degree of node u.

Here's some facts:

(1) "Stationary distribution" is $Pr[X = i] = \frac{d_i}{2m}$ ← (2 # edges $= \sum_{j=1}^{n} d_j$)

i.e. if $Pr[X_t = i] = \frac{d_i}{2m}$

(Q3 answer)

$\Rightarrow Pr[X_{t+1} = i] = Pr[X_t = i]$ (fixed point)

(2) Hitting time H_{uv}: expected time to get from u to v in random walk

(!)

We'll make connections to electrical networks to show:–

$\max_{uv} H_{uv} \le O(2m(n-1))$

(3) Cover time: C_u: expected time to see all vertices starting from u.

C_G: max ——— ——— from any vertex

$= \max_u C_u$.

Thm: $C_G \le 2m(n-1) \le O(n^3)$

These connections are more important than the actual theorems.

OK: Fact 1: sps. $p_i^{(t)} = \Pr[X_t = i] = \frac{d_i}{2m}$

then $p^{(t+1)} = p^{(t)}$

Pf: $p_i^{(t+1)} = \Pr[X_{t+1} = i] = \sum_j \Pr[X_{t+1} = i \mid X_t = j] \Pr[X_t = j]$

$= \sum_{j \sim i} p_j^{(t)} \cdot \frac{1}{d_j} = \sum_{j \sim i} \frac{1}{2m} = \frac{d_i}{2m}$

"j neighbor of i"

Another way to see it (the Laplacian connection)

$p^{(t)}$ row vector.

Claim: $p^{(t+1)} = p^{(t)} \cdot (D^{-1}A)$ $\quad D = \mathrm{diag}(d_1, \dots d_n)$

Pf: $(p_1^t \ \dots \ p_n^t) \begin{pmatrix} \frac{1}{d_1} & & & \\ & \frac{1}{d_2} & & \\ & & \ddots & \\ & & & \frac{1}{d_n} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & & \\ \vdots & & \\ A_{n1} & & \end{pmatrix}$

$= \left(\frac{p_1^t}{d_1} \ \cdots \ \frac{p_n^t}{d_n} \right) (A)$

$= \left(\sum_{j \to i} \frac{p_j^t}{d_j} \right)_i$

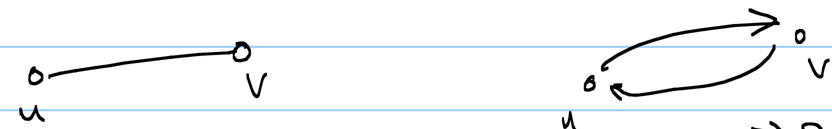
$\Rightarrow p$ is a stationary distribution if $p = p(D^{-1}A)$

$\Rightarrow p(I - D^{-1}A) = 0$

$\Rightarrow p \cdot D^{-1}(D - A) = 0 \Rightarrow (p \cdot D^{-1}) L = 0 \Rightarrow (p D^{-1}) = \alpha \mathbb{1}$ for some α $\Rightarrow p_i = \frac{d_i}{\sum d_i}$ ∎

<u>In fact</u>: Imagine the process being on vertex pairs u, v such that $\{u, v\} \in E$

Or bidirect each undirected edge



$\Rightarrow$ $2m$ directed edges. $\vec{E}$

<u>Claim</u>: walk on directed edges. If pick a random uniform one to start, then get uniformly-random one at next step

<u>Same Proof</u>: $\Pr[Y_t = (u, v)] = 1/2m$ say $\quad \forall (u,v) \in \vec{E}$

$$\text{then } \Pr[Y_{t+1} = (a, b)] = \sum_{c : (c,a) \in \vec{E}} \Pr[Y_t = (c,a)] \Pr[(a,b) | (c,a)]$$

$$= \sum_c \frac{1}{2m} \cdot \frac{1}{d_a} = d_a \cdot \frac{1}{2m} \cdot \frac{1}{d_a}$$

$$= 1/2m. \quad \blacksquare$$

So each edge equally likely (about $\frac{1}{2m}$ times on average)

. . . . . uv . . . . . . uv . . . . . uv.

And can use this (and connectivity of graph) to show that

$$\boxed{E[\text{time between successive visits to } \underbrace{(u,v)}_{\text{this edge}}] = 2m.}$$

Need to be a bit careful, see Avrim's notes on how to do this.

$\Rightarrow \quad H_{uv} + H_{vu} \leq 2m.$

Hand wavy, see Avrim's notes

"commute time"

We just saw intuition why $E[H_{uv} + H_{vu}] \leq 2m$ for edge (u,v)

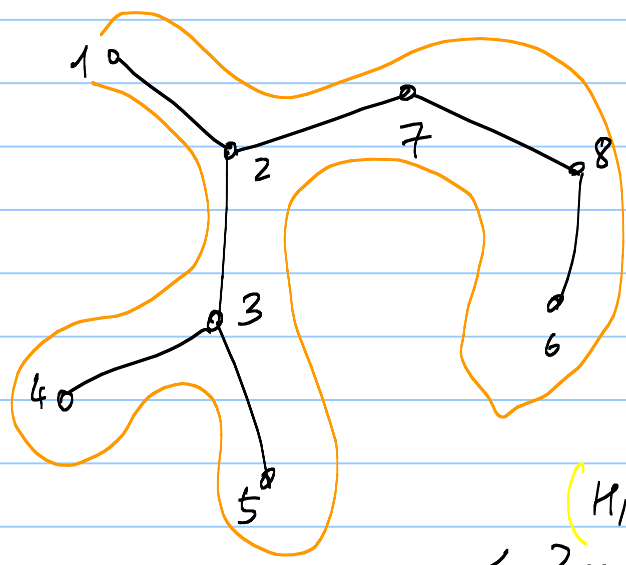
(will see another proof, more solid in a minute)

But for now: let's take this and show a fact about cover times

---

OK: says $E[\text{commute time}(u,v)] \leq 2m$. for any edge $(u,v) \in E$
$= H_{uv} + H_{vu}$

So: take G. fix a spanning tree on G



then cover time (G)

$$\leq H_{12} + H_{23} + H_{34} + H_{43} + H_{35} + H_{53} + H_{32} + \cdots\cdots + H_{87} + H_{72} + H_{21}$$

pair these up: —

$$(H_{12} + H_{21}) + (\cdot\ \cdot\ \cdot)\ \cdot\ \cdot$$

$$\leq \underbrace{2m + 2m + \cdots}_{n-1}$$

$$= 2m(n-1).$$

☺

So can explore any graph G in time $O(mn)$ without a map, only local information!

Pretty amazing.

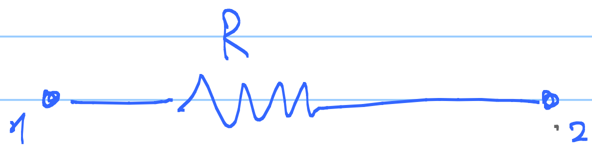
(Also, is tight — need this for some graphs)

For K_n (complete graph): Cover time $= \Theta(n \log n)$ much less.

For path: Cover time $= \Theta(n^2) = \Theta(mn)$.

Back to hitting times, lets get another proof using electrical networks!!

Remember Electrical Networks of Resistors?



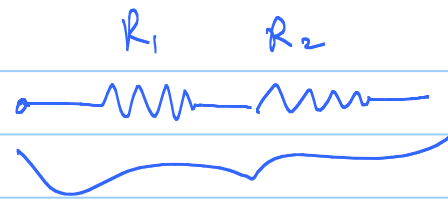
Put voltage ϕ_1, ϕ_2 on two nodes $\Rightarrow$ current $= \dfrac{\phi_1 - \phi_2}{R}$

And then there are laws:-

Ohm's Law

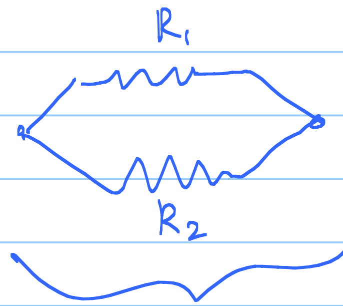
Kirchoff's Law: total current in = total current out

Can put resistors in series



$R_1 + R_2 = R_{eff}$ ← effective

or in parallel

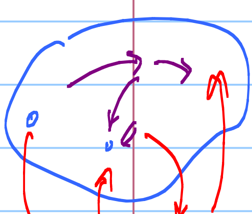


$\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{R_{eff}}$

(Fact: $R_{eff} \le R_i \ \forall i$ in parallel)

OK: Here's why things are useful.

We imagine our graph (undirected unweighted) as a network of resistors (edge = 1 Ω resistor)

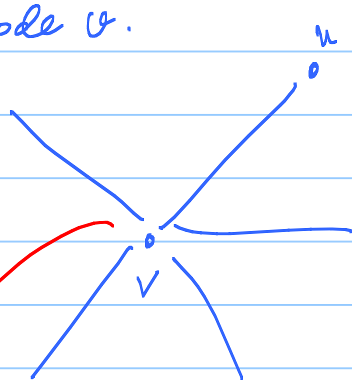


current flows inside network

current flowing to voltage source

Suppose we set some potentials (→ voltages) at nodes $\phi_1 \ldots \phi_n$
and have some current flow f on edges.

Look at node v.



flow on $\vec{uv} = \frac{\phi_u - \phi_v}{R_{uv}} = \phi_u - \phi_v$

So total current out of v

$= \sum_{u \sim v} (\phi_v - \phi_u)$

← current entering or leaving the network

$= d_v \phi_v - \sum_{u \sim v} \phi_u = (L \cdot \phi)_v$ ← Laplacian

$\Rightarrow \forall$ vertex v: $(L\phi)_v = b_v$ ← amount of current entering v from outside.

(by Kirchoff's Law)

<u>Good</u>: But also:— sps we fix a target node t and ask for expected hitting times to t. $h_x = E[H_{xt}]$

<u>then</u>: $\boxed{h_t = 0}$

$$h_x = 1 + \frac{1}{d_x}\sum_{y \sim x} h_y$$

$$\Updownarrow \quad d_x h_x = d_x + \sum_{y\sim x} h_y$$

$$\Leftrightarrow \quad d_x h_x - \sum_{y \sim x} h_y = d_x \Leftrightarrow \boxed{(Lh)_x = d_x. \quad \forall x \neq t}$$

Same equations! So can use one to solve for another.
(If they have unique solution)

In fact, just take the $n-1$ equations and solve for them:

same for current if you choose current at all nodes except t, current @ t fixed

- rank$(L) = n-1$ if graph is connected (since $n-1$ nonzero eigenvalues)

$\Rightarrow$ get $n-1$ independent inequalities, and hence solution.
may not have $h_t = 0$.

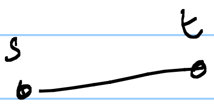
But can change all h values by same constant to set $h_t = 0$ since $\mathbb{1}$ is in kernel(L), does not change things.

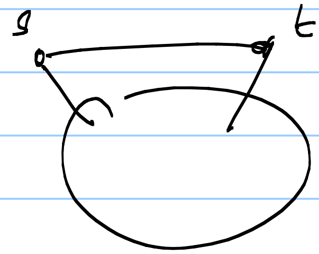
Theorem: Expected Commute time between s and t (for any s,t)

$$H_{st} + H_{ts} = 2m \cdot R_{eff}(s,t)$$

$R_{eff}(s,t)$ = effective resistance between s & t

= amount of voltage difference to send 1 unit of current from s to t.

Note:  if $(s,t) \in E \Rightarrow R_{eff}(s,t) \le 1$

because of parallel comb. 

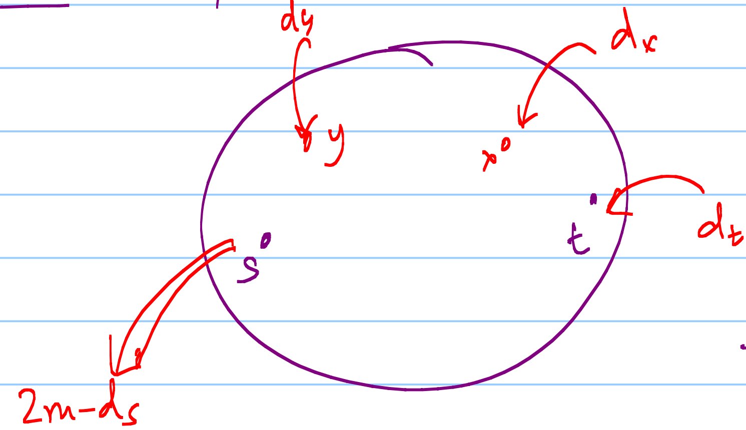
$\Rightarrow$ for an edge  in G, $H_{st} + H_{ts} \le 2m$.

Which is what we wanted to show.

Now to prove the theorem

Good: so push d_x flow in at all nodes except s

(⇒ take out $2m - d_s$ flow at s)

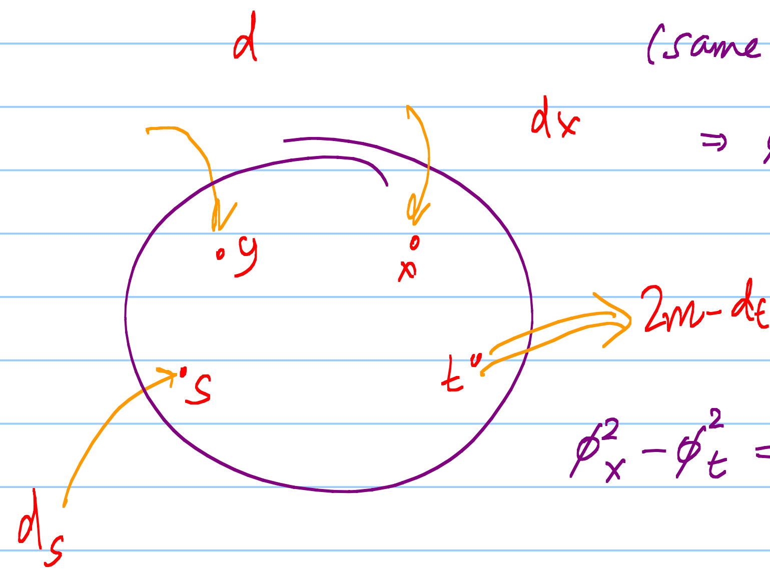


$$\Rightarrow (L\phi^1)_x = d_x \quad \forall x \neq s$$
$$\phi_s = 0$$

$$\phi^1_x - \phi^1_s = \text{Hitting time}_{xs} \quad \forall x$$

(same inequalities)

Similarly

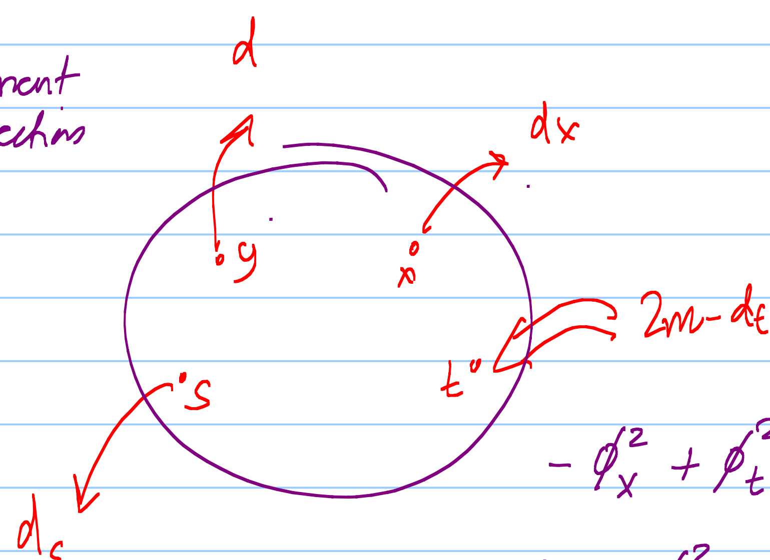


$$\Rightarrow \phi^1_t - \phi^1_s = H_{ts}$$

$$\phi^2_x - \phi^2_t = \text{Hitting time}_{xt} \quad \forall x$$

Flip the current directions

Only sign of voltages changes.



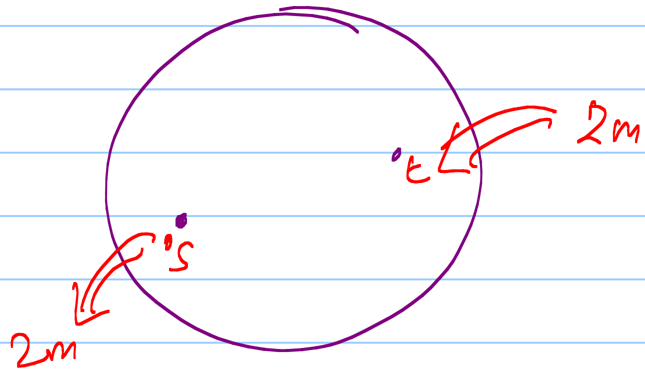
$$-\phi^2_x + \phi^2_t = \text{Hitting time}_{xt} \quad \forall x$$

$$\Rightarrow -\phi^2_s + \phi^2_t = H_{st}$$

Say set $\phi_s^1, \phi_s^2 = 0$

$$\Rightarrow \quad \phi_t^1 + \phi_t^2 = H_{st} + H_{ts}.$$

Now: combine the two electrical flows, potentials (add by linearity)



$$\phi_t^1 + \phi_t^2 - 0$$

potential of s

$$= H_{st} + H_{ts}.$$

this potential sends 2m current from s to t

$\Rightarrow$ potential = current × Reff.

$$\Rightarrow \quad 2m \cdot R_{eff}(s,t) = H_{st} + H_{ts}$$

∎

Other topics: pagerank.

: Markov Chains and rapid mixing (the eigenvalue connection)

: