TTIC 31150/CMSC 31150 Mathematical Toolkit (Spring 2023)

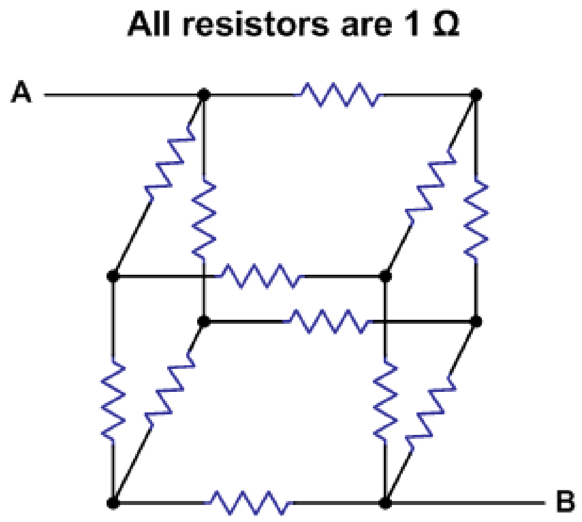
Avrim Blum and Ali Vakilian

Lecture 17: Random walks on graphs 2

Recap

- Random walks on undirected graphs: hitting time, commute time, cover time.
- Stationary distribution of random walk: uniform over edge/directions, or equivalently each node has probability proportional to its degree.
- **Theorem:** If G is a connected graph with n vertices and m edges, then $Cov_G \leq 2m(n-1)$.
- Electrical networks and connections to random walks.

Something completely different(?): electrical networks

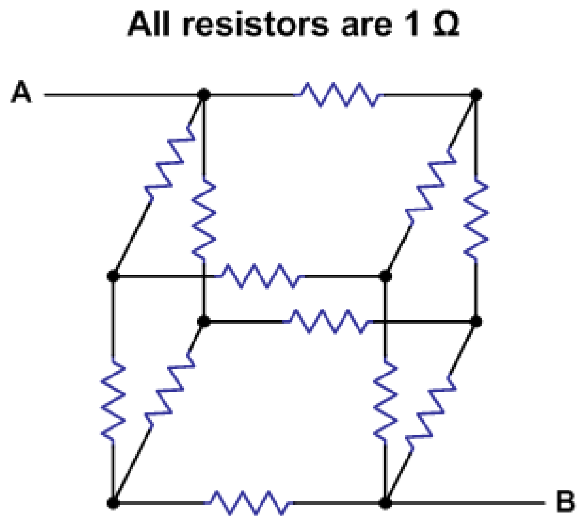


Consider a graph G where on each edge we have a resistor of some resistance.

- Say we connect a battery of some voltage V_{batt} between two nodes A and B (so $V_A - V_B = V_{batt}$, and let's for convenience say $V_B = 0$).
- Then each node in the graph will have a voltage (also called "potential") and each edge will have some current flowing in some direction.

Can think of voltage as like "height", and resistors like little water wheels or filters.

Something completely different(?): electrical networks



Voltages and currents can be computed using the following two rules.

- Kirchoff's law: current is like water flow: for any node not connected to the battery, flow in = flow out.
- Ohm's law: $V = IR$. Here, R is resistance, V is the voltage drop, and I is the current flow.

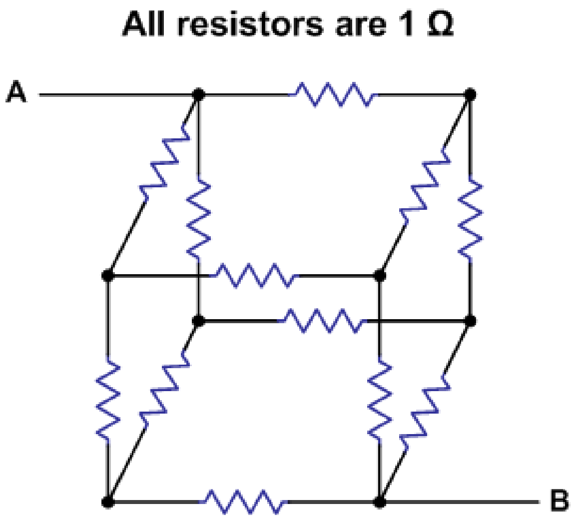
Effective resistance R_{uv} between u and v: connect up battery, measure current, $R_{uv} = \frac{V}{I}$.

Electrical networks and random walks

Consider a graph G, fix two distinguished nodes A,B.

Consider a random walk.

Let p_u be the probability a random walk starting from u reaches A before it reaches B.



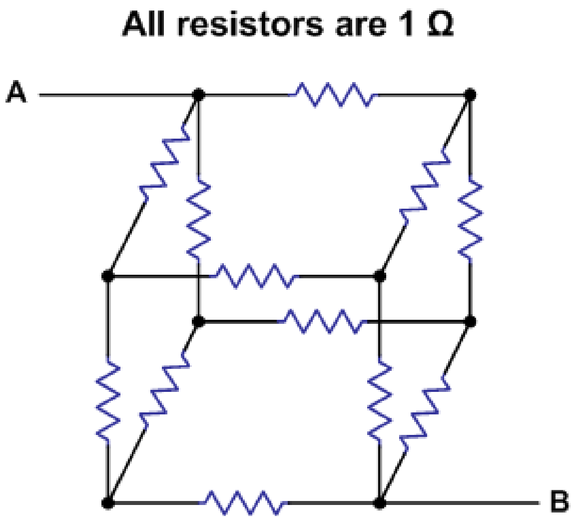
Consider placing a 1-volt battery between A and B

Let V_u be the voltage at node u.

Then $p_u = V_u$.

- Solving for p_u: $p_A = 1$, $p_B = 0$, and for all $u \notin \{A, B\}$ we have $p_u = \frac{1}{\deg(u)} \sum_{v:\{u,v\}\in E} p_v$.
- Solving for V_u: $V_A = 1$, $V_B = 0$, and for all $u \notin \{A, B\}$ we have flow in = flow out, which means $V_u = \frac{1}{\deg(u)} \sum_{v:\{u,v\}\in E} V_v$.

Another connection: effective resistance and commute time

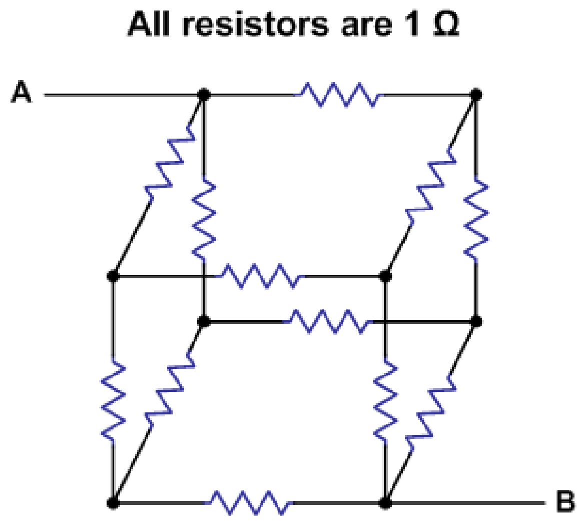


Theorem: In a connected graph G with m edges, each of which is a unit resistor, for any two nodes u, v we have $C_{uv} = 2mR_{uv}$.

- For example, on a line graph of n nodes and $n-1$ edges, the commute time between the two endpoints is exactly $2(n-1)^2$.
- Note that if u, v are neighbors then $R_{uv} \leq 1$, so $C_{uv} \leq 2m$. (So, this is another proof of the main lemma from last time).

Another connection: effective resistance and commute time

Example computation of effective resistance



Theorem: In a connected graph G with m edges, each of which is a unit resistor, for any two nodes u, v we have $C_{uv} = 2mR_{uv}$.

Key lemma

Lemma: Fix some vertex v. For each node $x \neq v$, place battery of voltage H_{xv} with positive terminal at x and negative terminal at v. Then $\deg(x)$ current will flow out of each $x \neq v$ and $2m - \deg(v)$ current will flow into v.

Proof:

- Let's define v to have voltage 0, so each node x has voltage H_{xv} $(H_{vv} = 0)$.
- For $x \neq v$, by definition of hitting time: $H_{xv} = 1 + \frac{1}{\deg(x)} \sum_{w:\{x,w\}\in E} H_{wv}$
- Current on edge (x, w) is $(V_x - V_w)/1$. So, total current flowing out of $x \neq v$ is:

$$\sum_{w:\{x,w\}\in E} V_x - V_w = \sum_{w:\{x,w\}\in E} H_{xv} - H_{wv} = \deg(x) \cdot H_{xv} - \sum_{w:\{x,w\}\in E} H_{wv} = \deg(x).$$

- And so $2m - \deg(v)$ current is flowing into v.

Key lemma #2

Lemma: Fix some vertex v. For each node $x \neq v$, place battery of voltage H_{xv} with positive terminal at x and negative terminal at v. Then $\deg(x)$ current will flow out of each $x \neq v$ and $2m - \deg(v)$ current will flow into v.

Lemma: Fix some vertex u. For each node $x \neq u$, place battery of voltage H_{xu} with negative terminal at x and positive terminal at u. Then $\deg(x)$ current will flow into each $x \neq u$ and $2m - \deg(u)$ current will flow out of u.

Proof: Same (or by symmetry: if you reverse all the batteries, you reverse all the currents).

Now, let's prove the theorem from the two lemmas.

Proof of theorem from lemmas

Lemma: Fix some vertex v. For each node $x \neq v$, place battery of voltage H_{xv} with positive terminal at x and negative terminal at v. Then $\deg(x)$ current will flow out of each $x \neq v$ and $2m - \deg(v)$ current will flow into v.

Lemma: Fix some vertex u. For each node $x \neq u$, place battery of voltage H_{xu} with negative terminal at x and positive terminal at u. Then $\deg(x)$ current will flow into each $x \neq u$ and $2m - \deg(u)$ current will flow out of u.

- Consider adding the voltages from the two experiments. So, voltage drop from u to v of $H_{uv} + H_{vu} = C_{uv}$.
- If add voltages, then currents add too by linearity. This gives us $2m$ units of current flowing out of u and $2m$ flowing into v.
- Since no current flowing into/out of any other node, can view as just a battery between u and v.
- Using $V = IR$ we get $C_{uv} = 2m \cdot R_{uv}$.

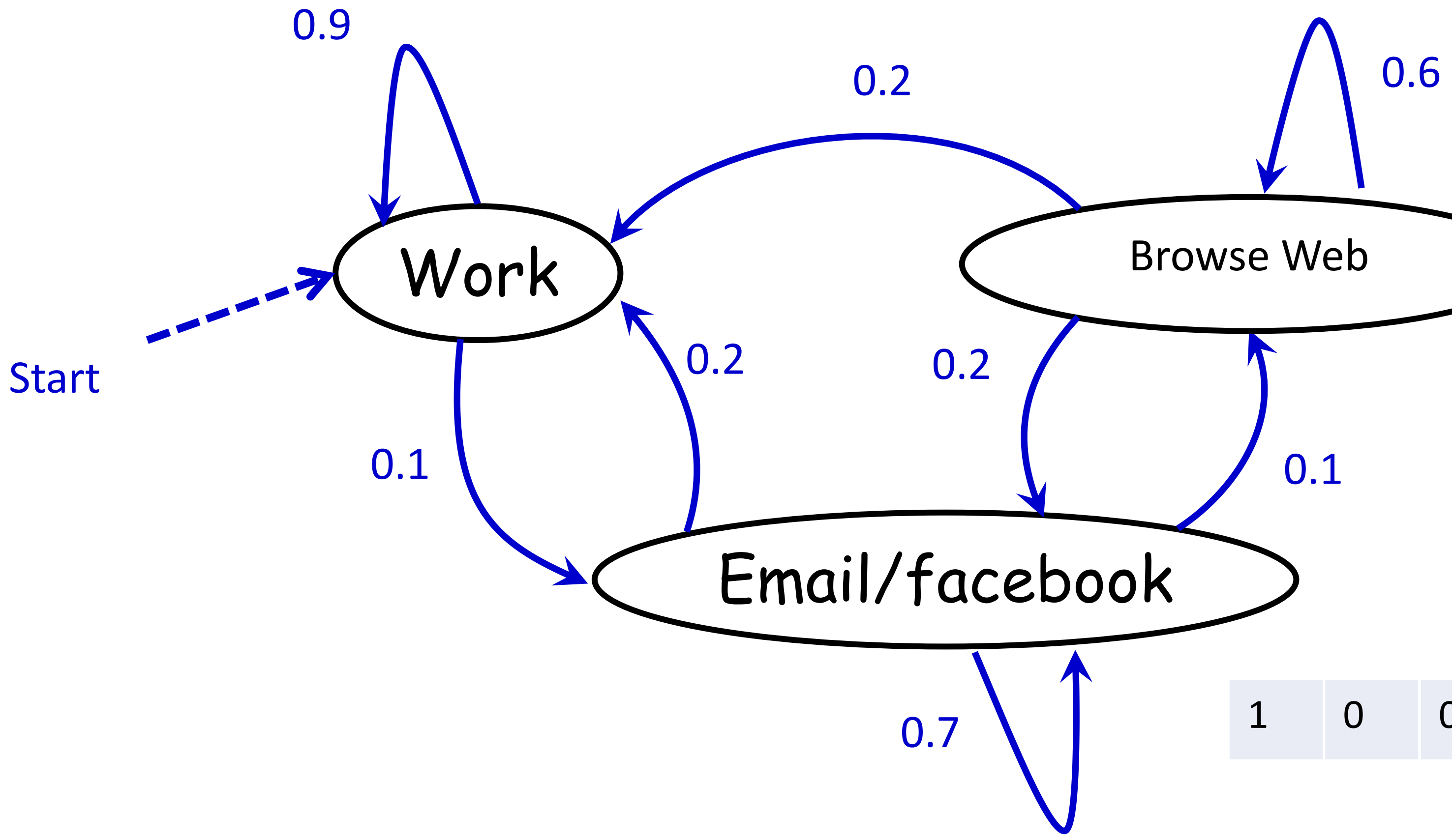
Markov Chains

A **Markov Chain** can be thought of as a random walk on a weighted directed graph:

- n states.
- An $n \times n$ transition matrix P where P_{ij} is the probability of moving to state j given that you currently are in state i.
- If you describe your current state as a row vector q then your next state is qP.
- Often used to describe probabilistic processes.

Markov Chain Example

Say you are planning to work on your homework but are easily distracted:



	Work	Email	Web
Work	0.9	0.1	0
Email	0.2	0.7	0.1
Web	0.2	0.2	0.6

1	0	0

0.9	0.1	0
0.2	0.7	0.1
0.2	0.2	0.6

=

0.9	0.1	0

More definitions

A **Markov Chain** can be thought of as a random walk on a weighted directed graph:

- n states.
- An $n \times n$ transition matrix P where P_{ij} is the probability of moving to state j given that you currently are in state i.
- If you describe your current state as a row vector q then your next state is qP.
- If underlying graph (directed edges with nonzero probability) is strongly connected, then it's *irreducible*.
- Irreducible Markov Chain is *aperiodic* if for every start state q there exists some T such that qP^T has nonzero probability on every state.

For example, a random walk on a complete bipartite graph would be irreducible but not aperiodic. If you add self-loops, then it becomes aperiodic.

More definitions

A **Markov Chain** can be thought of as a random walk on a weighted directed graph:

- n states.
- An $n \times n$ transition matrix P where P_{ij} is the probability of moving to state j given that you currently are in state i.
- If you describe your current state as a row vector q then your next state is qP.
- If underlying graph (directed edges with nonzero probability) is strongly connected, then it's *irreducible*.
- Irreducible Markov Chain is *aperiodic* if for every start state q there exists some T such that qP^T has nonzero probability on every state.
- A stationary distribution π is a left eigenvector of eigenvalue 1. That is, $\pi = \pi P$.

Stationary distributions

- A stationary distribution π is a left eigenvector of eigenvalue 1. That is, $\pi = \pi P$.

- This is the largest eigenvalue, because for any vector v (even if it has negative entries), the sum of absolute values cannot increase when multiplying by P. I.e., $\|v\|_1 \geq \|vP\|_1$.

- Because every row P_i sums to 1, so $|v_i| = \sum_j |v_i P_{ij}|$. So, $\|vP\|_1 \leq \sum_i \|v_i P_i\|_1 = \|v\|_1$.

Symmetric Markov chains

A Markov chain is symmetric if P is symmetric. E.g., a random walk on an undirected graph where every node has the same degree.

- For a symmetric Markov chain, all column sums are 1, so the stationary distribution is uniform. [“The” stationary distribution if the MC is connected, else “a” stationary distribution if not]

- One way to see it: columns summing to one and $\pi = \pi P$ means that each π_i is a weighted average of the others. [can you see the rest of the proof?]

Rapid Mixing

Often we will want to define a Markov chain on a "solution space" whose size is exponential in the natural problem parameters. E.g., each state could be an assignment of values to n variables.

In this case, we have no hope to visit the entire state space, but perhaps we can more quickly approach the stationary distribution?

A Markov chain is rapidly mixing if can get close to stationary in $polylog(n)$ steps.

Example: random walk on the cube $\{0,1\}^d$. Here $n = 2^d$. To make this aperiodic, let's say that at each step we stay put with probability ½.

- Equivalent walk: at each step, pick a random coordinate, replace with uniform random 0/1 value.

Rapid Mixing

For example, a symmetric MC

Theorem 2.1 *Say P is a Markov chain with real eigenvalues and orthogonal eigenvectors. Then, for any starting distribution $q^{(0)}$, the L_2 distance between the distribution after T steps $q^{(T)} = q^{(0)}P^T$ and the stationary distribution π is at most $|\lambda_2|^T$ where λ_2 is the eigenvalue of largest absolute value among eigenvectors orthogonal to π.*

- So, if $|\lambda_2| \leq 1 - \epsilon$, then for any constant c it takes only $T = O\left(\frac{\log n}{\epsilon}\right)$ steps to get $\left\|q^{(T)} - \pi\right\|_2 \leq 1/n^c$.
- What happened to irreducibility and aperiodicity? If reducible or periodic, then $|\lambda_2| = 1$ so theorem is vacuous. E.g., complete bipartite graph has eigenvector with all nodes on the left assigned $1/n$ and all nodes on the right assigned $-1/n$ with eigenvalue -1.

Rapid Mixing

For example, a symmetric MC

Theorem 2.1 *Say P is a Markov chain with real eigenvalues and orthogonal eigenvectors. Then, for any starting distribution $q^{(0)}$, the L_2 distance between the distribution after T steps $q^{(T)} = q^{(0)}P^T$ and the stationary distribution π is at most $|\lambda_2|^T$ where λ_2 is the eigenvalue of largest absolute value among eigenvectors orthogonal to π.*

Proof:

- Let's say the orthogonal eigenvectors are $v_1, \dots, v_n$ with $v_1 = \pi$.
- They form a basis, so can write $q^{(0)} = c_1\pi + c_2v_2 + c_3v_3 + \cdots + c_nv_n$ for some $c_1, \dots, c_n$.
- After T steps, we have $q^{(T)} = c_1\pi + c_2\lambda_2^Tv_2 + c_3\lambda_3^Tv_3 + \cdots + c_n\lambda_n^Tv_n$.
- Assuming $|\lambda_2| < 1$ (else the theorem is vacuously true) note that this approaches $c_1\pi$ as $T \to \infty$. This means we must have $c_1 = 1$.

Rapid Mixing

For example, a symmetric MC

Theorem 2.1 *Say P is a Markov chain with real eigenvalues and orthogonal eigenvectors. Then, for any starting distribution $q^{(0)}$, the L_2 distance between the distribution after T steps $q^{(T)} = q^{(0)}P^T$ and the stationary distribution π is at most $|\lambda_2|^T$ where λ_2 is the eigenvalue of largest absolute value among eigenvectors orthogonal to π.*

Proof:

- Let's say the orthogonal eigenvectors are $v_1, \ldots, v_n$ with $v_1 = \pi$.
- They form a basis, so can write $q^{(0)} = c_1\pi + c_2v_2 + c_3v_3 + \cdots + c_nv_n$ for some $c_1, \ldots, c_n$.
- After T steps, we have $q^{(T)} = c_1\pi + c_2\lambda_2^Tv_2 + c_3\lambda_3^Tv_3 + \cdots + c_n\lambda_n^Tv_n$.
- So, $\left\|q^{(T)} - \pi\right\|_2 = \left\|c_2\lambda_2^Tv_2 + \cdots + c_n\lambda_n^Tv_n\right\|_2 \leq |\lambda_2|^T \cdot \|c_2v_2 + \cdots + c_nv_n\|_2 \leq |\lambda_2|^T$.

By orthogonality

Since $\left\|q^{(0)}\right\|_2 \leq \left\|q^{(0)}\right\|_1 = 1$

That’s it….

- Final exam will be made available on Monday.
- Can download and take it when you like: you have 24 hours to turn it in from the time you download the exam. Turn it in via dropbox link.
- All exams should be turned in by 11:59pm Friday night May 26 (11:59pm Thursday night if you are graduating this quarter)
- Please also fill in the course evals – we read them all and they are useful to us in improving the course.