Outline

- One Concrete Example of a random walk
- Motivation applications
  - shuffling cards
  - universal traverse sequence
  - self stabilizing token management scheme
  - random sampling
  - enumeration and volume computation
- Definitions
  - random walk
  - markov chain
  - stationary walk
- Main Parameters
  - access time
  - commute time
  - cover time
  - mixing rate
- examples path, cycle, complete graph lollipop, expanders
- bounds on the main parameters
- spectra connection
- efficient stopping rules OR reversal and forget time
- spanning trees
Two concrete examples

Example of a random walk on $P_6$. Starting at 2, find the probability of reaching 3 for the first time in three steps.

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}.$$

Example of a random walk on $P_3$. Starting at 1, find the hitting time for 3. That is, the expected number of steps to reach 3 for the first time.

$$2 \cdot \frac{1}{2} + 3 \cdot 0 + 4 \cdot \frac{1}{4} + 5 \cdot 0 + 6 \cdot \frac{1}{8} + \cdots = \sum_{i=1}^{\infty} \frac{2i}{2^i} = \sum_{i=1}^{\infty} \frac{i}{2^{i-1}} \leq 4.525$$

References


Mark Jerrum, *Counting, Sampling and Integrating: Algorithms and Complexity*

http://www.dcs.ed.ac.uk/home/mrj/pubs.html

(scroll down that page)

Motivating Applications

How many shuffles?

Goal: Reach the uniform distribution on the sample space of all possible orderings on a deck of $k$ cards.
Digraph Model:
   Vertices - $k!$ permutations of the cards.
   Edges - from $u$ to $v$ if there exists a riffle shuffle from $u$ to $v$.

Find $x$ so that given any two vertices $u$ and $v$, no matter where we start, after $x$ shuffles (a random walk of length $x$) the probability that we end at $u$ equals the probability that we end at $v$.

Universal Traverse Sequence (Computer Science)

Let $G = (V, E)$ be a connected $d$-regular graph and $v_0 \in V$. Suppose at each vertex, there is a label associated with each of its $d$ edges. Labels $1, 2, \ldots, d$.

A traverse sequence is a sequence $(h_1, h_2, \ldots, h_t) \in \{1, 2, \ldots, d\}^t$ such that starting at $v_0$ take edge labeled $h_1$ from that vertex take edge labeled $h_2$, and so on. Then we visit all nodes. A universal traverse sequence with parameters $n$ and $d$ is a traverse sequence that works for every $d$-regular graph on $n$ vertices given any labeling and starting point.

**Theorem 1** (Aleliunas, Karp, Lipton, Lovasz, Packoff) For any $d \geq 2$ and $n \geq 3$, there exists a universal traverse sequence of length $O(d^2 n^3 \log n)$.

Notice that choosing one element at random from $\{1, 2, \ldots, d\}^t$ is the same as a random walk, given any given starting point.

Self stabilizing token management scheme for a computer network

The token is the authorization for the processor carrying it to perform some task. Only one processor is supposed to have it at any one time. Suppose, by some disturbance, two processors carry it. They pass it around randomly until the two tokens collide. From then on, the system is back to normal. How long does this take?

Let $M(u, v)$ denote the expected number of steps before two random walks starting at vertices $u$ and $v$ collide.

Let $H(u, v)$ be the expected number of steps a random walk starting at $u$ takes to reach $v$ for the first time.
Theorem 2 (Coppersmith, Tetali, Winkler)

\[ M(u, v) \leq H(u, v) + H(v, w) - H(w, u) \]

for some vertex \( w \).

Random walks can be used to generate random elements in large and complicated sets.

As in card shuffling.

Also, selecting a perfect matching from all perfect matchings of a given graph. Worst case, the complete graph on \( n \) vertices has

\[ \binom{n}{2} \cdot \binom{n-2}{2} \cdot \binom{n-4}{2} \cdots \]

different perfect matchings.

Define a graph - with perfect matchings as vertices and edges between 2 that differ by switching 2 edges. Begin a random walk at \( \{\{1, 2\}, \{3, 4\}, \ldots, \{n-1, n\}\} \) and be sure to go long enough, \( t \) steps, so that the distribution of ending points is close to the uniform distribution. This can be done in a polynomial in \( n \) steps.


Enumeration and Volume Computation

Suppose I want to estimate the total number of facts about opera. Suppose I was going to be in a contest where I was going to be asked questions about opera. I know zero about it. So I go and memorize 10 facts. Then you ask me 6 questions at random. I know the answer to only 1 of them. Let \( T \) = the total number of facts about opera. Then

\[ \frac{10}{T} \approx \frac{1}{6}, \quad T \approx 60. \]

We can apply this method to counting the number of perfect matchings in bipartite graphs as long as we have a way to generate random ones.
Definitions

Random Walk. Given a graph and a starting vertex we select a neighbor, uniformly at random and move to this neighbor. Then select another neighbor at random, move to it, etc. The random sequence of vertices is a random walk.

We will restrict ourselves to finite undirected connected graphs, but there are other examples.

- Brownian motion of a dust particle in a room. $B(t)$ is the position in $\mathbb{R}^3$ at time $t \in \mathbb{R}$.

- Grid graphs - the $d$-dimensional lattice $L_d$. Polya (1921) if we do a random walk on $L_d$ we return to the starting point infinitely often if $d = 2$ and only a finite number of times if $d \geq 3$.

- Digraphs - used to model a nonsymmetric relationship, like shuffling cards.

More formal

Let $G = (V, E)$. Consider a random walk on $G$. Start at a vertex $v_0$ at the $t^{th}$ step we are at vertex $v_t$ move to a neighbor of $v_t$ with probability $1/\text{deg}(v_t)$.

The sequence of random vertices $(v_0, v_1, \ldots,)$ is a finite markov chain that is time-reversible, due to the fact that the underlying graph is undirected.

Review a little probability

A random process $X$ is a collection of random variables $\{X_t : t \in T\}$ taking on values in $S$, state space. If $T = \{0, 1, 2, \ldots\}$, we call it a discrete-time process. Usually the random variables are dependent on one another as in random evolution or the random walk.

For a random walk, we have a random process $X$. Let $X_i = v_i$. Notice that if we are at $v_i$ in the $i^{th}$ step, the probability that $X_{i+1} = v_{i+1}$ does not depend on the values $X_0, X_1, \ldots, X_{i-1}$. But it does depend on the value of $X_i$.

Defn 1 The process $X$ is a markov chain if it satisfies the markov condition

$$\text{Prob}(X_n = s | X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1})$$
\[ = \text{Prob}(X_n = s | X_{n-1} = x_{n-1}) \]
for all \( n \geq 1 \) and all \( s, x_0, x_1, \ldots x_{n-1} \in S \).

Review a little probability

**Defn 2** The chain is called **homogeneous** if

\[ \text{Prob}(X_{n+1} = j | X_n = i) = \text{Prob}(X_1 = j | X_0 = i) \]
for all \( n, i, j \). The transition matrix \( P = (p_{i,j}) \) is the \(|S| \times |S|\) matrix of transition probabilities

\[ p_{i,j} = \text{Prob}(X_{n+1} = j | X_n = i). \]

Random Walks on graphs

Let \((v_0, v_1, \ldots, v_t)\) be a random walk on a graph \( G = (V, E) \)
Let \( v_0 \) be selected randomly according to an initial distribution \( P_0 \).
Let \( P_s \) be the distribution of \( v_s \), so that \( P_s(i) = \text{Prob}(v_s = i) \).
The transition matrix of the markov chain is given by

\[ M = (p_{i,j}), \quad p_{i,j} = \begin{cases} \frac{1}{\text{deg}(i)} & \{i, j\} \in E \\ 0 & \text{otherwise} \end{cases} \]

Let \( A \) be the adjacency matrix of \( G \) and \( D \) the diagonal matrix with \( D_{i,i} = 1/\text{deg}(v_i) \). Then \( M = DA \). (\( M^t \) is the transpose of \( M \).)

Rule of Walk: \( P_{s+1} = (M^t)P_s = (M^t)*P_0 \).

Example

Let \( G = (V, E) \) with \( V = \{0, 1, 2, 3, 4\} \) and

\[ E = \{\{0, 1\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}\} \]
**Example**

Let $P_0 = \left( \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5} \right)$.

\[
M = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
1/3 & 0 & 1/3 & 0 & 1/3 \\
0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 1/2 & 0 & 1/2 \\
0 & 1/2 & 0 & 1/2 & 0
\end{bmatrix}
\]

\[
P_1 = M^t P_0 = \left[ .0667, .4000, .1667, .2000, .1667 \right]
\]

\[
P_2 = M^t P_1 = \left[ .1333, .2333, .2333, .1667, .2333 \right]
\]

\[
P_{10} = (M^t)^{10} P_0 = \left[ .1200, .2400, .2400, .1600, .2400 \right]
\]

\[
P_{20} = (M^t)^{20} P_0 = \left[ .1200, .2400, .2400, .1600, .2400 \right]
\]

**The stationary distribution**

**Defn 3** Consider the sequence of distributions:

\[(P_0, P_1, P_2, \ldots)\]

We say $P_0$ is stationary if $P_0 = P_1$ which then implies $P_t = P_0, \forall t \geq 0$. This is called the stationary walk.

**Fact 1** For every $G$, $\pi(v) = \frac{\text{deg}(v)}{2m}$ is stationary, where $m$ is the number of edges.
**Fact 2** Let $G$ be connected and not bipartite. The stationary distribution is unique.

**Fact 3** If $G$ is connected and not bipartite, $P_t \to \Pi$ as $t \to \infty$, for any starting distribution $P_0$.

**Proof of fact 1**

\[
M'\pi(u) = \sum_{uv \in E} \frac{1}{\deg(v)} \cdot \frac{\deg(v)}{2m} \\
= \frac{1}{2m} \sum_{u,v \in E} 1 \\
= \frac{\deg(u)}{2m} \\
= \pi(u)
\]

**Looking back at previous example**

Let $P_0 = (.1, .3, .2, .2, .2)$.

\[
M = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
1/3 & 0 & 1/3 & 0 & 1/3 \\
0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 1/2 & 0 & 1/2 \\
0 & 1/2 & 0 & 1/2 & 0
\end{bmatrix}
\]

We get

\[
P_1 = M'P_0 = [.1, .3, .2, .2, .2]
\]

**Stochastic Matrices**

**Defn 4** An $n \times n$ matrix $A$ is **Stochastic** if $A$ has nonnegative entries and the row sums equal 1.
Fact 4 A has a positive eigenvector, namely the all ones vector \([1,1,1,\ldots,1]\). The corresponding eigenvalue, 1, is \(\rho(A)\). Where \(\rho(A)\) is the maximum absolute value over all eigenvalues of \(A\).

For above fact, see Cor 8.1.30 in *Matrix Analysis* by Horn and Johnson.

If \(A\) is also positive then the following hold, see Thm 8.2.4 and 8.2.5 in *Matrix Analysis* by Horn and Johnson.

Fact 5 \(A\) has no other eigenvalues of modulus 1 other than 1 itself.

Fact 6 The eigenspace belonging to the eigenvalue 1 has geometric multiplicity 1.

Stochastic Matrices

Fact 7 Suppose \(A\) is stochastic with strictly positive diagonal entries Let \(x\) be an eigenvector belonging to the eigenvalue 1 of \(A^t\). Then for every \(x_0\), the Markov-chain sequence defined by

\[x_0, x_1 = A^t x_0, x_2 = A^t x_1 = (A^t)^2 x_0, \ldots, x_i = A^t x_{i-1} = (A^t)^i x_0, \ldots\]

converges to a scalar multiple of \(x\). The value of the scalar is \(\alpha = \text{the sum of the entries of} \ \pi\).

See *Applied Linear Algebra*, Cor 9.35 by Nobel and Daniel.

Proof of Fact 2: Unique stationary dist

Let \(G\) be connected and not bipartite. There exists an \(s\) such that \((M^t)^s\) has all positive entries, \(s = \text{diam}(G)\). See the maple worksheet for an example.

Notice that we don’t get this if \(G\) is bipartite.

In fact, \([(M^t)^s]_{i,j} = p_{i,j}^s\) which is the probability that a walk starting at \(j\) reaches \(i\) in \(s\) steps.

Notice that the column sums of \((M^t)^s\) are for column \(i\):
\[
\sum_{j=1}^{n} \text{Prob}(\text{a walk starting at } i \text{ reaches } j \text{ in } t \text{ steps}) = 1
\]

By Fact 1, \( \pi \) is the stationary distribution of \((M^t)^s\) which implies, \( \pi = (M^t)^s \pi \). So \( \pi \) is an eigenvector belonging to eigenvalue 1. By Fact 6, this is unique (up to scalar multiples.) But \( \pi \) is the unique scalar multiple of itself that represents a probability distribution.

**Proof of fact 3:**

Let \( G \) be connected and not bipartite. There exists an \( s \) such that \((M^t)^s\) has all positive entries, \( s = \text{diam}(G) \). Then both Fact 2 and Fact 7 hold.

We wish to show, \( P_r \to \Pi \) as \( r \to \infty \), for any starting distribution \( P_0 \), where \( P_0, P_1, P_2, \ldots, P_r, \ldots \) is given below.

\[
P_0, P_1 = (M^t)P_0, P_2 = (M^t)^2P_0, \ldots, P_r = (M^t)^rP_0, \ldots, \tag{1}
\]

From Fact 7, for \( \pi \), the eigenvector belonging to the eigenvalue 1, the Markov-chain sequence defined by

\[
x_0, x_1 = (M^t)x_0, x_2 = ((M^t)^s)x_0, \ldots, x_i = ((M^t)^s)^i x_0, \ldots
\]

converges to a scalar multiple of \( \pi \), for any \( x_0 \). The value of the scalar is \( \alpha = \) the sum of the entries of \( \pi \).

Replacing \( x_0 \) with \((M^t)^{s-1}P_0\), we get

\[
P_{s-1} = (M^t)^{s-1}P_0, P_{2s-1} = (M^t)^{2s-1}P_0, P_{3s-1} = (M^t)^{3s-1}P_0, \ldots,
\]

\[
\ldots, P_{(i+1)s-1} = (M^t)^{(i+1)s-1}P_0, \ldots
\]

converges to \( \alpha \pi \).

If we started with \((M^t)^{s+1}\) which is also positive, we get,

\[
P_s = (M^t)^sP_0, P_{2s} = (M^t)^{2s}P_0, P_{3s} = (M^t)^{3s}P_0, \ldots,
\]

\[
\ldots, P_{(i+1)s} = (M^t)^{(i+1)s}P_0, \ldots
\]

converges to \( \alpha \pi \).
Similarly, we could start with \((M^t)^{s+\ell}\) for any \(ell \in [s-1]\) to get
\[
P_{s+\ell-1} = (M^t)^{s+\ell-1}P_0; P_{2s+\ell-1} = (M^t)^{2s+\ell-1}P_0; P_{3s+\ell-1} = (M^t)^{3s+\ell-1}P_0; \ldots;
\]
\[
\ldots, P_{(i+1)s+\ell-1} = (M^t)^{(i+1)s+\ell-1}P_0; \ldots.
\]
converges to \(\alpha \pi\). Thus the whole sequence, \(P_0, P_1, P_2, \ldots, P_r, \ldots\) converges to \(\alpha \pi\).

**Time reversible**

A random walk considered backwards is still a random walk. In fact, consider all walks \((v_0, v_1, v_2, \ldots, v_t)\) resulting from the initial distribution \(P_0\) and ending in \(P_t\). This will be the same as the set of walks resulting from the initial distribution \(P_t\). It won’t necessarily end in \(P_0\).

Using the stationary distribution, this means \(\pi(i)p_{i,j} = \pi(j)p_{j,i}\).

Note: Markov chains are said to be time reversible under the same condition as above.

**Review Expectations**

If \(X\) is a numeric random variable taking on possible values in a discrete sample space, \(S\), then
\[
EX(X) = \sum_{s \in S} s \text{Prob}(X = s).
\]

**Example 0.1** Say a random walk just passed through vertex \(u\). What is the expected number of steps before returning to \(u\)? ANS: (LL) \(2m/\text{deg}(u)\).

**Example 0.2** Say you just passed through edge \(vu\) from \(v\) to \(u\), what is the expected number of steps before returning to \(u\) by the same edge? ANS: (LL) \(2m\).

**Example 0.3** Let \(n_t(x)\) denote the number of times \(x\) occurs in \(v_0, v_1, \ldots, v_{t-1}\). Then \(EX(n_t(x)/t) \rightarrow \text{deg}(x)/2m\) as \(t \rightarrow \infty\). So that, for \(N\) large enough, a random walk with \(N\) steps passes through \(v\) about \(N \cdot \text{deg}(v)/2m\) times. (LL)
Main Parameters

- **Access time** or **Hitting time**. \( H_{i,j} \) is the expected number of steps before \( j \) is visited starting from vertex \( i \). The sum \( K(i,j) = H_{i,j} + H_{j,i} \) is called the **commute time** from \( i \) to \( j \).

- The **commute time** is the expected number of steps to reach every node. If no starting vertex is given, we mean the worst case.

- The **mixing rate** is the measure of how fast the random walk converges to its limiting distribution. If \( G \) is non-bipartite then \( p^{(s)}_{i,j} \) is the \( i,j \)-entry of \((M^t)^s\). Then \( p^{(s)}_{i,j} \rightarrow \deg(j)/2m \) as \( s \rightarrow \infty \). See Fact 3. The mixing rate is defined as

\[
\mu = \lim_{t \to \infty} \sup_{i,j} \max_{s} |p^{(s)}_{i,j} - \deg(j)/2m|^{1/s}.
\]

Some examples - Path

The path \( P_{n+1} \) on \( n+1 \) vertices, \( \{0,1,\ldots,n\} \). Find \( H_{i,j} \).

Start with \( H_{n,n} \), the expected number of steps starting at \( n \) until returning to \( n \). Let \( X \) be the number of edges used. Then

\[
X = \sum_{e \in E} \text{number of times } e \text{ is crossed}.
\]

\[EX(X) = \text{expected number of times } \{n-1,n\} \text{ is used} + \sum_{e \notin \{n-1,n\}} \text{expected number of times } e \text{ is used} = 2 + H'_{n-1,n-1}
\]

where \( H'_{n-1,n-1} \) is the access time from \( n-1 \) to \( n-1 \) in \( P_n \). This recursive formula gives \( 2n \).

Now we find \( H_{n-1,n} \): Since from vertex \( n \), we are forced to move to \( n-1 \) first, we see that \( H_{n,n} = H_{n-1,n} + 1 \). So \( H_{n-1,n} = 2n - 1 \).
Examples - Path - continued

Now try $H_{k-1,k}$, $1 \leq k \leq n$. $H_{k-1,k} = 2k - 1$ since the vertices to the right of $k$ play no role.

Now find $H_{i,k}$, $1 \leq i < k \leq n$. We see that

$$H_{i,k} = H_{i,k-1} + H_{k-1,k}$$
$$= H_{i,k-1} + 2k - 1$$
$$= H_{i,i+1} + H_{i+1,i+2} + \cdots + H_{k-1,k}$$
$$= 2(i + 1) - 1 + 2(i + 2) - 1 + \cdots + 2k - 1$$
$$= \sum_{j=i+1}^{k} (2j - 1) = 2 \sum_{j=i+1}^{k} j - \sum_{j=i+1}^{k} 1$$
$$= 2 \left[ \binom{k + 1}{2} - \binom{i + 1}{2} \right] - (k - i)$$
$$= (k + 1)k - (i + 1)i - k + i = k^2 - i^2$$

Summary of Path

$H_{n,n} = 2n$. $H_{n-1,n} = 2n - 1$. $H_{k-1,k} = 2k - 1$. $H_{i,k} = k^2 - i^2$, for $0 \leq i < k \leq n$. $H_{0,0} = H_{n,n}$. $H_{k,i} = H_{n-k,n-i}$, for $0 \leq i < k \leq n$.

The cover time, starting with 0 is $n^2$

What is the cover time, starting with $k$?