Random Walks on Graphs

Josep Fàbrega

Journées Combinatoire et Algorithmes du Littoral Mediterranéen

UPC, Barcelona, 7–8 June 2011

Outline

Simple random walks on graphs
- Random walks and Markov chains
- Mixing rate. Hitting, commute and cover times

Random walks and harmonic functions
- Connection with electrical networks

Random walks on weighted graphs
- Recurrence in infinite graphs

Random walks in algorithm design

References

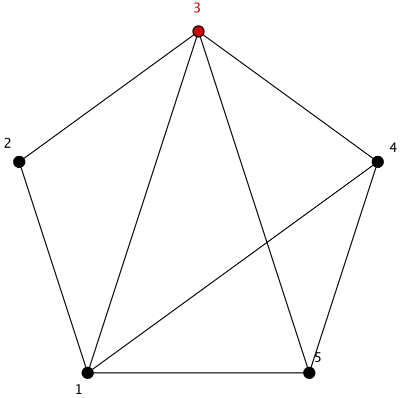
- Doyle, P.G. and Snell, J. L., Random walks and electric networks, Carus Mathematical Monographs vol. 22 (1984).
- Grimmett, G., Probability on graphs, Cambridge Univ. Press (2010).
- Lovász, L., Random walks on graphs: a survey, Combinatorics, Paul Erdős is eighty, Vol. 2, Keszthely (1993).

Simple random walks

Let $G = (V, E)$ be a connected graph, $|V| = n$ and $|E| = m$.

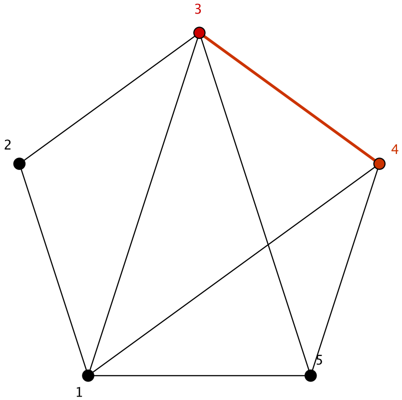
- Given an initial vertex v_0, select "at random" an adjacent vertex v_1, and move to this neighbour.
- Then select "at random" a neighbor v_2 of v_1, and move to it; etc.

Simple random walks



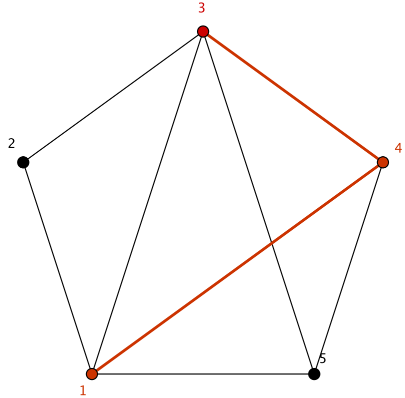
$$v_0 = 3$$

Simple random walks



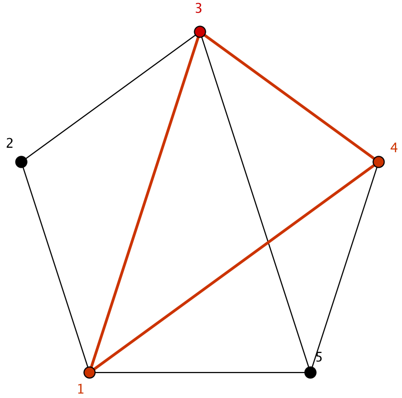
$3 \to 4$

Simple random walks



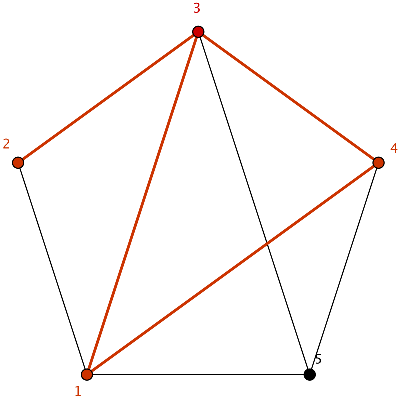
$3 \to 4 \to 1$

Simple random walks



$$3 \to 4 \to 1 \to 3$$

Simple random walks



$$3 \to 4 \to 1 \to 3 \to 2 \to \cdots$$

Simple random walks

The sequence of vertices $v_0, v_1, v_2, \ldots, v_k, \ldots$ selected in this way is a simple random walk on G.

At each step k, we have a random variable X_k taking values on V.

Hence, the random sequence

$$X_0, X_1, X_2, \ldots, X_k, \ldots$$

is a discrete time stochastic process defined on the state space V.

Simple random walks

What does "at random" mean?

If at time k we are at vertex i, choose uniformly an adjacent vertex $j \in \Gamma(i)$ to move to.

Let $d(i)$ denote the degree of vertex i.

$$p_{ij} = \mathrm{P}(X_{k+1} = j \mid X_k = i) = \begin{cases} \dfrac{1}{d(i)}, & \textit{if } ij \in E \\ 0, & \textit{otherwise} \end{cases}$$

These transition probabilities do not depend on "time" k.

Simple random walks

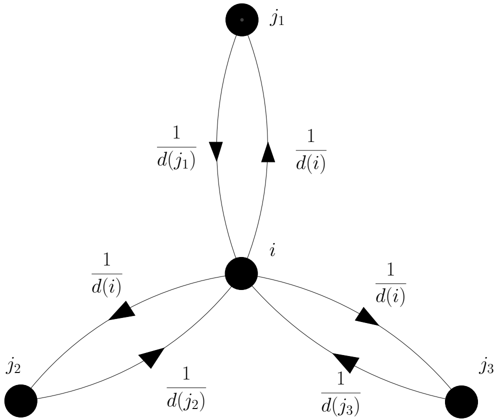


Figure: Transition probabilities

Random walks and Markov chains

The Markov property holds: conditional on the present, the future is independent of the past.

$$\begin{aligned}\mathrm{P}(X_{k+1} = j \mid X_k = i, X_{k-1} = i_{k-1}, \ldots, X_0 = i_0)&\\ = \mathrm{P}(X_{k+1} = j \mid X_k = i) = p_{ij}&\end{aligned}$$

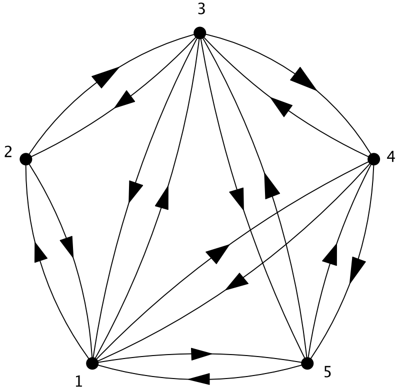
The random sequence of vertices visited by the walk,

$$X_0, X_1, \ldots, X_n, \ldots$$

is a Markov chain with state space V and matrix of transition probabilities

$$\mathbf{P} = (p_{ij})_{i,j \in V}$$

Random walks and Markov chains



$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

Random walks and Markov chains

$\mathbf{P}$ is a stochastic matrix,

$$\sum_{j \in V} p_{ij} = 1$$

Let $\mathbf{D}$ be the diagonal matrix with $(\mathbf{D})_{ii} = 1/d(i)$ and $\mathbf{A}$ be the adjacency matrix of G. Then

$$\mathbf{P} = \mathbf{D}\,\mathbf{A}$$

In particular, if G is d-regular,

$$\mathbf{P} = \frac{1}{d}\,\mathbf{A}$$

Random walks and Markov chains

Let $\boldsymbol{\rho}_k$ be the row vector giving the probability distribution of X_k,

$$\boldsymbol{\rho}_k(i) = \mathrm{P}(X_k = i), \quad i \in V$$

The rule of the walk is expressed by the simple equation

$$\boldsymbol{\rho}_{k+1} = \boldsymbol{\rho}_k \, \mathbf{P}$$

That is, if $\boldsymbol{\rho}_0$ is the initial distribution from which the starting vertex v_0 is drawn,

$$\boldsymbol{\rho}_k = \boldsymbol{\rho}_0 \, \mathbf{P}^k, \quad k \geq 0$$

The stationary distribution

Since G is connected, the random walk on G corresponds to an irreducible Markov chain.

The Perron-Frobenius theorem for nonnegative matrices implies the existence of a unique probability distribution $\boldsymbol{\pi}$, which is a positive left eigenvector of $\mathbf{P}$ associated to its dominant eigenvalue $\lambda = 1$.

$$\boldsymbol{\pi} = \boldsymbol{\pi}\,\mathbf{P},$$

$$\boldsymbol{\pi}(i) > 0 \quad \textit{for all } i \in V, \quad \sum_{i \in V} \boldsymbol{\pi}(i) = 1$$

The stationary distribution

If the initial vertex of the walk is drawn from $\boldsymbol{\pi}$, then the probability distribution at time k is

$$\boldsymbol{\rho}_k = \boldsymbol{\pi}\,\mathbf{P}^k = \boldsymbol{\pi}$$

Hence, for all time $k \geq 0$,

$$\mathrm{P}(X_k = i) = \boldsymbol{\pi}(i), \quad i \in V$$

The random walk is a stationary stochastic process.

Detailed balance condition

If we find a probability distribution $\boldsymbol{\pi}$ *which satisfies the detailed balance condition,*

$$\boldsymbol{\pi}(i)\, p_{ij} = \boldsymbol{\pi}(j)\, p_{ji}, \quad \text{for all } i, j \in V,$$

then it is the stationary distribution.

Certainly, if this condition is satisfied, we have for all $i \in V$,

$$\boldsymbol{\pi}(i) = \boldsymbol{\pi}(i) \sum_{j \in V} p_{ij} = \sum_{j \in V} \boldsymbol{\pi}(i)\, p_{ij} = \sum_{j \in V} \boldsymbol{\pi}(j)\, p_{ji}$$

that is,

$$\boldsymbol{\pi} = \boldsymbol{\pi}\, \mathbf{P}$$

Detailed balance condition

For a random walk on G, it is straightforward to find a probability vector satisfying the detailed balance condition,

$$\frac{\boldsymbol{\pi}(i)}{d(i)} = \frac{\boldsymbol{\pi}(j)}{d(j)} = k, \quad ij \in E$$

But

$$1 = \sum_{i \in j} \boldsymbol{\pi}(i) = k \sum_{i \in V} d(i) = k\,2m$$

Hence,

$$k = \frac{1}{2m}$$

Detailed balance condition

Therefore,

The stationary probabilities are proportional to the degrees of the vertices,

$$\boldsymbol{\pi}(i) = \frac{d(i)}{2m}, \quad \textit{for all } i \in V$$

In particular, if G is d-regular,

$$\boldsymbol{\pi}(i) = \frac{d}{2m} = \frac{1}{n}, \quad \text{for all } i \in V$$

and $\boldsymbol{\pi}$ is the uniform distribution.

Time-reversibility

The detailed balance condition implies time-reversibility.

Suppose that the random walk has the stationary distribution and consider the reversed walk

$$Y_n = X_{m-n}, \quad n = 0, 1, \ldots, m$$

Then,

$$\begin{aligned} &\mathrm{P}(Y_{n+1} = j \mid Y_n = i) \\ &\quad = \frac{\mathrm{P}(Y_n = i \mid Y_{n+1} = j)\, \boldsymbol{\pi}(j)}{\boldsymbol{\pi}(i)} = \frac{p_{ji}\, \boldsymbol{\pi}(j)}{\boldsymbol{\pi}(i)} = p_{ij} \end{aligned}$$

The reversed walk is also a Markov chain and looks the same as X.

Expected return times

The property of time-reversibility means

$$\pi(i)\, p_{ij} = \pi(j)\, p_{ji}$$

Hence, a stationary walk steps as often from i to j as from j to i.

Moreover,

$$\pi(i)\, p_{ij} = \frac{1}{2m} \quad \text{for all } ij \in E$$

A random walk moves along every edge, in every given direction, with the same frequency $1/2m$.

If the random walk just passed through an edge, then the expected number of steps before it traverses again the same edge in the same direction is $2m$.

Expected return times

A similar fact holds for vertices:

If the random walk just visited vertex i, then the expected number of steps before it returns to i is

$$\frac{1}{\pi(i)} = \frac{2m}{d(i)}$$

When G is regular this return time is just n, the number of nodes.

Convergence to the limiting distribution

When G is non-bipartite, the Markov chain is aperiodic.

In this case, the stationary distribution is also a limiting distribution.

If G is non-bipartite, then for any $i, j \in V$

$$\mathrm{P}(X_k = j \mid X_0 = i) \to \boldsymbol{\pi}(j), \quad \textit{as } k \to \infty,$$

The convergence to $\boldsymbol{\pi}(j)$ does not depend on the initial vertex i.

Convergence to the limiting distribution

Hence,

$$\mathsf{P}(X_k = j) = \sum_{i \in V} \mathsf{P}(X_k = j \mid X_0 = i)\,\mathsf{P}(X_0 = i) \to \boldsymbol{\pi}(j),$$

that is,

$$\boldsymbol{\rho}_k \to \boldsymbol{\pi},$$

independently of the initial distribution.

This is equivalent to

$$\mathbf{P}^k \to \boldsymbol{\Pi},$$

where $\boldsymbol{\Pi}$ is a stochastic matrix with all its rows equal to $\boldsymbol{\pi}$.

Convergence to the limiting distribution

To prove this result, let us bring $\mathbf{P}$ to a symmetric form. Recall that $\mathbf{P} = \mathbf{D}\,\mathbf{A}$. Consider the symmetric matrix

$$\mathbf{N} = \mathbf{D}^{1/2}\mathbf{A}\mathbf{D}^{1/2} = \mathbf{D}^{-1/2}\mathbf{P}\mathbf{D}^{1/2}$$

which has the same eigenvalues as $\mathbf{P}$,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

Write $\mathbf{N}$ in spectral form

$$\mathbf{N} = \sum_{r=1}^{n} \lambda_r \mathbf{v}_r^T \mathbf{v}_r$$

(The row eigenvectors $\mathbf{v}_r$ are unitary and orthogonal.)

Convergence to the limiting distribution

It is easily checked that $\mathbf{w} = \left(\sqrt{d(1)}, \ldots, \sqrt{d(n)}\right)$ is a positive eigenvector of $\mathbf{N}$ with eigenvalue 1.

It follows from the Perron-Frobenius Theorem that

$$\lambda_1 = 1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$$

Moreover, if G is non-bipartite $\lambda_n > -1$.

Notice that $\mathbf{v}_1$ must be vector $\mathbf{w}$ normalized

$$\mathbf{v}_1 = \frac{1}{\sqrt{2m}} \mathbf{w} = \left(\sqrt{\pi(1)}, \ldots, \sqrt{\pi(n)}\right)$$

Convergence to the limiting distribution

We have

$$\mathbf{P}^k = \mathbf{D}^{1/2}\,\mathbf{N}^k\,\mathbf{D}^{-1/2} = \sum_{r=1}^{n} \lambda_r^k\,\mathbf{D}^{1/2}\,\mathbf{v}_r^T\mathbf{v}_r\,\mathbf{D}^{-1/2}$$

Notice that

$$\left(\mathbf{D}^{1/2}\,\mathbf{v}_1^T\mathbf{v}_1\,\mathbf{D}^{-1/2}\right)_{ij} = \frac{\sqrt{\pi(i)}}{\sqrt{d(i)}}\,\sqrt{\pi(j)}\,\sqrt{d(j)} = \frac{d(j)}{2m} = \pi(j)$$

Thus,

$$\mathbf{D}^{1/2}\,\mathbf{v}_1^T\mathbf{v}_1\,\mathbf{D}^{-1/2} = \mathbf{\Pi}$$

Then,

$$\mathbf{P}^k = \mathbf{\Pi} + \sum_{r=2}^{n} \lambda_r^k\,\mathbf{D}^{1/2}\,\mathbf{v}_r^T\mathbf{v}_r\,\mathbf{D}^{-1/2}$$

Convergence to the limiting distribution

Hence,

$$\left(\mathbf{P}^k\right)_{ij} = \boldsymbol{\pi}(j) + \sum_{r=2}^{n} \lambda_r^k \, v_{ki} v_{kj} \sqrt{\frac{d(j)}{d(i)}}$$

If G is non-bipartite, $|\lambda_r| < 1$ for $r = 2, \ldots, n$. Therefore, as $k \to \infty$,

$$\mathrm{P}(X_k = j \mid X_0 = i) = \left(\mathbf{P}^k\right)_{ij} \to \boldsymbol{\pi}(j)$$

Or, for any initial distribution,

$$\mathrm{P}(X_k = j) \to \boldsymbol{\pi}(j)$$

Convergence to the limiting distribution

Let

$$\lambda = \max\{|\lambda_2|, |\lambda_n|\}$$

Theorem

For a random walk starting at node i

$$|\mathrm{P}(X_k = j) - \boldsymbol{\pi}(j)| \leq \sqrt{\frac{d(j)}{d(i)}}\, \lambda^k$$

When G is non-bipartite the convergence to the stationary distribution is geometric with ratio λ.

Mixing rate

The mixing rate is a measure of how fast the random walk converges to its limiting distribution,

$$\mu = \limsup_{k\to\infty} \max_{i,j} \left| (\mathbf{P}^k)_{ij} - \frac{d(j)}{2m} \right|^{1/k}$$

The number of steps before the distribution of X_k will be close to the limiting distribution is about $\log n/(1-\mu)$.

This mixing time may be much less than the number of nodes. For an expander graph, this takes only $\mathcal{O}(\log n)$ steps.

Mixing rate

From

$$|\mathrm{P}(X_k = j) - \boldsymbol{\pi}(j)| \leq \sqrt{\frac{d(j)}{d(i)}}\,\lambda^k$$

we see that the mixing rate is at most λ.

Indeed, equality holds.

Theorem

The mixing rate of a random walk on a non-bipartite graph is

$$\lambda = \max\{|\lambda_2|, |\lambda_n|\}$$

Hitting and commute times

The access time or hitting time $H(i, j)$ is the expected number of steps before node j is visited, starting from node i.

In general

$$H(i, j) \neq H(j, i)$$

The commute time is

$$\kappa(i, j) = H(i, j) + H(j, i)$$

$\kappa(i, j)$ is a symmetric parameter.

Access times for a path

Let us determine the access time for 2 vertices of a path on nodes $0, 1, \ldots, n$

$H(k-1, k)$ is one less than the expected return time of a random walk on a path of length k, starting at the last node.

$$H(k-1, k) = 2k - 1$$

Now, consider $H(i, k)$, $0 \leq i < k \leq n$.

In order to reach k, we have to reach node $k - 1$. From here, we have to get to k.

$$H(i, k) = H(i, k - 1) + 2k - 1$$

Access times for a path

Hence,

$$\begin{aligned} H(i,k) &= H(i,k-1) + 2k - 1 \\ &= H(i,k-2) + (2k-3) + (2k-1) \\ &= \cdots \cdots \\ &= H(i,i+1) + (2i+3) + \cdots + (2k-1) \\ &= (2i+1) + (2i+3) + \cdots + (2k-1) = k^2 - i^2 \end{aligned}$$

In particular,

$$H(0,k) = k^2$$

Symmetry properties of the access times

If a graph has a vertex-transitive automorphism group, then $H(i,j) = H(j,i)$ for all nodes i and j.

Coppersmith, Tetali and Winkler (1993): For any three nodes i, j and k,

$$H(i,j) + H(j,k) + H(k,i) = H(i,k) + H(k,j) + H(j,i)$$

The nodes of any graph can be ordered so that if i precedes j then $H(i,j) \leq H(j,i)$.

A hitting probability

Let T_i be the first time when a random walk starting at i returns to i and T_{ij} the first time when it returns to i after visiting j.

Observe that $T_i \leq T_{ij}$.

Let

$$p = \mathrm{P}(T_i = T_{ij})$$

be the probability that a random walk starting at i visits j before returning to i.

Notice that

$$\mathrm{E}(T_i) = \frac{1}{\boldsymbol{\pi}(i)} = \frac{2m}{d(i)}, \qquad \mathrm{E}(T_{ij}) = \kappa(i,j)$$

A hitting probability

We have

$$
\begin{aligned}
\mathsf{E}(T_{ij}) - \mathsf{E}(T_i) &= \mathsf{E}(T_{ij} - T_i) \\
&= p\,\mathsf{E}(T_{ij} - T_i \mid T_i = T_{ij}) + (1-p)\,\mathsf{E}(T_{ij} - T_i \mid T_i < T_{ij}) \\
&= (1-p)\,\mathsf{E}(T_{ij}) = \mathsf{E}(T_{ij}) - p\,\mathsf{E}(T_{ij})
\end{aligned}
$$

Hence,

$$
p = \frac{\mathsf{E}(T_i)}{\mathsf{E}(T_{ij})} = \frac{2m}{d(i)\,\kappa(i,j)}
$$

Theorem

The probability that a random walk starting at i visits j before returning to i is

$$
\frac{2m}{d(i)\,\kappa(i,j)}
$$

Access times and spectra

A firs-step analysis gives, for $i \neq j$,

$$H(i,j) = \sum_{k \in \Gamma(i)} (1 + H(k,j))\, p_{ik} = 1 + \frac{1}{d(i)} \sum_{k \in \Gamma(i)} H(k,j)$$

and, for any $i \in V$,

$$\frac{2m}{d(i)} = 1 + \frac{1}{d(i)} \sum_{k \in \Gamma(i)} H(k,i)$$

Access times and spectra

These two equations can be put together in matrix notation.

Let $\mathbf{H}$ be a square matrix such that $(\mathbf{H})_{ij} = H(i,j)$ and $(\mathbf{H})_{ii} = 0$.

Then,

$$\mathbf{H} + 2m\,\mathbf{D} = \mathbf{J} + \mathbf{P}\mathbf{H}$$

That is,

$$(\mathbf{I} - \mathbf{P})\,\mathbf{H} = \mathbf{J} - 2m\,\mathbf{D}$$

We can not solve for $\mathbf{H}$ because $\mathbf{I} - \mathbf{P}$ is singular.

Access times and spectra

Let

$$\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{\Pi})^{-1}$$

It is easily checked that

$$\mathbf{H} = \mathbf{J} - 2m\,\mathbf{Z}\,\mathbf{D} + \mathbf{\Pi}\,\mathbf{H}$$

Hence,

$$\begin{aligned} H(i,j) &= 1 - \frac{2m}{d(j)}(\mathbf{Z})_{ij} + (\boldsymbol{\pi}\mathbf{H})_j \\ 0 &= 1 - \frac{2m}{d(j)}(\mathbf{Z})_{jj} + (\boldsymbol{\pi}\mathbf{H})_j \end{aligned}$$

Access times and spectra

Thus, we can compute the access times from the fundamental matrix $\mathbf{Z}$:

$$H(i,j) = 2m\,\frac{(\mathbf{Z})_{jj} - (\mathbf{Z})_{ij}}{d(j)}$$

Diagonalizing $\mathbf{N} = \mathbf{D}^{-1/2}\mathbf{P}\mathbf{D}^{1/2}$ as above we get

Theorem

$$H(i,j) = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{kj}^2}{d(j)} - \frac{v_{ki}v_{kj}}{\sqrt{d(i)d(j)}}\right)$$

Access times and spectra

Two corollaries are

Corollary

$$\kappa(i,j) = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}}{\sqrt{d(j)}} - \frac{v_{ki}}{\sqrt{d(i)}} \right)^2$$

$$m \left(\frac{1}{d(i)} + \frac{1}{d(j)} \right) \leq \kappa(i,j) \leq \frac{2m}{1-\lambda_2} \left(\frac{1}{d(i)} + \frac{1}{d(j)} \right)$$

The difference $1 - \lambda_2$ is called the spectral gap.

Cover time

The cover time (starting from a given distribution) is the expected number of steps to reach every node.

For instance, if we start from an endnode, the cover time of the path of length n is

$$H(0, n) = n^2$$

(It suffices to reach the other endnode.)

Cover time for a complete graph

Consider a complete graph on nodes $\{0, \ldots, n-1\}$.

If we start at 0, the probability that we first reach vertex 1 in the k-th step is

$$\left(\frac{n-2}{n-1}\right)^{k-1} \frac{1}{n-1}$$

This is a geometric distribution with parameter $1/(n-1)$. Hence, the expected time to reach 1 is

$$H(0,1) = n-1$$

(Of course, $H(i,j) = H(0,1)$ for any $i \neq j$.)

Cover time for a complete graph

Now, let T_i be the first time when i vertices have been visited,

$$T_1 = 0 < T_2 = 1 < T_3 < \ldots < T_n$$

Notice that $T_{i+1} - T_i$ is the number of steps while we wait for a new vertex. This is an event with probability $(n-i)/(n-1)$.

Hence,

$$\mathrm{E}(T_{i+1} - T_i) = \frac{n-1}{n-i}$$

Thus, the cover time is

$$\mathrm{E}(T_n) = \sum_{i=1}^{n-1} \mathrm{E}(T_{i+1} - T_i) = \sum_{i=1}^{n-1} \frac{n-1}{n-i} \approx n \log n$$

Bounds for the cover time

It is conjectured that the graph with smallest cover time is the complete graph.

Aldous (1989): There exists a constant $c > 0$ such that for every graph with n vertices, the cover time τ_n satisfies

$$\tau_n \geq c\, n \log n$$

provided that the starting vertex is selected at random from the stationary distribution.

Bounds for the cover time

Feige (1995): The cover time τ_n from any starting node in a graph with n nodes satisfies

$$\tau_n \geq (1 - o(1))\; n \log n$$

$$\tau_n \leq \left(\frac{4}{27} + o(1) \right)\; n^3$$

Feige (1993): For a regular graph on n vertices,

$$\tau_n \leq 2n^2$$

Why $\tau_n \leq \mathcal{O}(n^3)$?

Let $ij \in E$. A firs-step analysis gives

$$\frac{2m}{d(i)} = \frac{1}{d(i)} \sum_{k \in \Gamma(i)} (1 + H(k,i))$$

Hence,

$$2m = \sum_{k \in \Gamma(i)} (1 + H(k,i)) \geq 1 + H(j,i)$$

Therefore, the access time between the endvertices i, j of an edge satisfies

$$H(j,i) < 2m$$

Why $\tau_n \leq \mathcal{O}(n^3)$?

Now, consider a spanning tree $\mathcal{T}$ with a distinguished vertex v.

The number of steps to traverse $\mathcal{T}$ (starting from v), and so covering all the vertices of G is at most $2n$.

Now, start from v a random walk on G. To walk (in G) from one endvertex of an edge of $\mathcal{T}$ to its other endvertex, the expected number of steps is at most $2m$.

Therefore,

The cover time starting from any vertex v is at most $4nm$. Hence,

$$\tau_n \leq 4nm$$

Random walks and harmonic functions

Let $G = (V, E)$ be a connected graph and $S \subseteq V$.

A function $\phi : V \to \mathbb{R}$ is a harmonic function with boundary S if

$$\frac{1}{d(i)} \sum_{j \in \Gamma(i)} \phi(j) = \phi(i)$$

holds for every $i \in V \setminus S$

Random walks and fair games

For a random walk, a harmonic function has the following interpretation.

Suppose at a given time k the random walk is visiting vertex i.

$$
\begin{aligned}
&\mathrm{E}(\phi(X_{k+1}) \mid X_k = i) \\
&\qquad = \sum_{j \in V} \phi(j)\, p_{ij} = \frac{1}{d(i)} \sum_{j \in \Gamma(i)} \phi(j) = \phi(i)
\end{aligned}
$$

Thus, the stochastic process

$$\phi(X_0), \phi(X_1), \ldots, \phi(X_k), \ldots$$

is a martingale with respect to X. (We are playing a fair game.)

Example: hitting probabilities

Let $S = \{s, t\}$.

Let $\phi(i)$ denote the probability that a random walk starting at i hits s before it hits t.

By conditioning on the first step we have for every $i \in V \setminus S$

$$\phi(i) = \sum_{j \in V} \phi(j)\, p_{ij} = \frac{1}{d(i)} \sum_{j \in \Gamma(i)} \phi(j)$$

Also,

$$\phi(s) = 1, \quad \phi(t) = 0$$

That is, ϕ is harmonic with boundary $\{s, t\}$.

Example: hitting probabilities

More generally,

Let $S \subseteq V$ and suppose we have a function $\phi_0 : S \longrightarrow \mathbb{R}$.

Let $\phi(i)$ be the expected valued of $\phi_0(s)$, where s is the random vertex where the random walk started at i first hits S.

Again

$$\phi(i) = \frac{1}{d(i)} \sum_{j \in \Gamma(i)} \phi(j)$$

and ϕ is harmonic with boundary S.

Random walks and electrical networks

Consider the graph G as an electrical network, where each edge represents a unit resistance.

Suppose that an electric current is flowing through G, entering at s and leaving at t.

This current flow is described by Kirchhoff laws.

Let $\phi(i)$ be the voltage of node i and f_{ij} be the current flowing from node i to an adjacent node j. Ohm's law implies

$$f_{ij} = \phi(i) - \phi(j)$$

Random walks and electrical networks

Kirchhoff's current law states that the total current flowing out of any vertex $i \in V \setminus \{s, t\}$ is zero.

$$0 = \sum_{j \in \Gamma(i)} f_{ij} = \sum_{j \in \Gamma(i)} (\phi(i) - \phi(j)) = d(i)\phi(i) - \sum_{j \in \Gamma(i)} \phi(j)$$

Again,

$$\phi(i) = \frac{1}{d(i)} \sum_{j \in \Gamma(i)} \phi(j)$$

and ϕ is a harmonic function with boundary $\{s, t\}$.

Harmonic functions provide an interesting connection between random walks on graphs and electrical networks.

Properties of harmonic functions

$\phi(i)$ lies between the minimum and maximum of ϕ over S.

Given $S \subseteq V$ and $\phi_0 : S \longrightarrow \mathbb{R}$, there is a unique harmonic function on G with boundary S extending ϕ_0.

- existence: follows by construction.
- uniqueness: consider the maximum of the difference of two such functions.

Uniqueness implies that the functions ϕ of the previous examples (the hitting probabilities for the random walk and the potential in the electric network) are the same.

Commute time and effective resistance

Let $\phi(i)$ be the potential of node i when we put a current through G from s to t, where $\phi(s) = 1$ and $\phi(t) = 0$.

The total current entering G by s is equal to the total current leaving it by t

$$\sum_{i \in \Gamma(t)} \phi(i)$$

Thus, let

$$R_{st} = \left(\sum_{i \in \Gamma(t)} \phi(i) \right)^{-1}$$

be the effective resistance between nodes s and t.

Commute time and effective resistance

On the other hand, $\phi(i)$ is the probability that a random walk starting at i visits s before t.

Therefore, the probability that a random walk starting at t hits s before returning to t is

$$\frac{1}{d(t)} \sum_{i \in \Gamma(t)} \phi(i) = \frac{1}{d(t) R_{st}}$$

But this probability is also equal to

$$\frac{2m}{d(t)\, \kappa(s,t)}$$

Commute time and effective resistance

Equating the two expressions,

Theorem

Consider G as an electrical network and let R_{st} denote the effective resistance between nodes s and t. Then the commute time between s and t is

$$\kappa(s,t) = 2m\, R_{st}$$

Adding any edge to G does not increase any resistance R_{st}.

Thus, by adding and edge no commute time is increased by more than a factor $(m+1)/m$.

Commute time and spanning trees

Using topological formulas from the theory of electrical networks, we get

Corollary

Let G' be the graph obtained from G by identifying s and t, and let $\tau(G)$ denote the number of spanning trees of G. Then

$$\kappa(s,t) = 2m\,\frac{\tau(G')}{\tau(G)}$$

Random walks on weighted graphs

Random walks can be generalized to graphs with weighted edges:

$$\begin{aligned} w : E &\longrightarrow (0, \infty) \\ e = ij &\mapsto w_{ij} \end{aligned}$$

Let $w(i) = \sum_{j \in \Gamma(i)} w_{ij}$.

Now, the transition probabilities are

$$p_{ij} = \mathrm{P}(X_{k+1} = j \mid X_k = i) = \begin{cases} \dfrac{w_{ij}}{w(i)}, & \textit{if } ij \in E \\ 0, & \textit{otherwise} \end{cases}$$

Random walks on weighted graphs

The detailed balance condition implies

$$\boldsymbol{\pi}(i)\,\frac{w_{ij}}{w(i)} = \boldsymbol{\pi}(j)\,\frac{w_{ji}}{w(j)}$$

But, $w_{ij} = w_{ji}$. Hence,

$$\frac{\boldsymbol{\pi}(i)}{w(i)} = k, \quad ij \in E$$

Random walks on weighted graphs

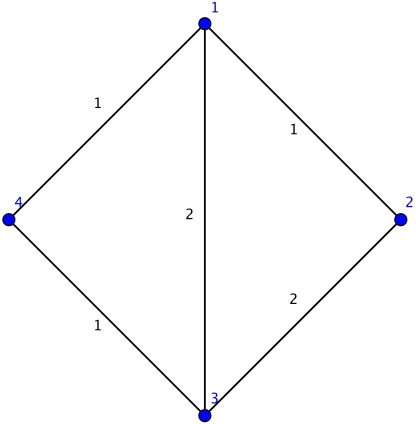
Since $\sum_{i \in V} \pi(i) = 1$, we have $k = 1/W$, where $W = \sum_{i \in V} w(i)$.

The stationary distribution is given by

$$\boldsymbol{\pi}(i) = \frac{w(i)}{W}, \quad \textit{for all } i \in V$$

In the regular case, $w(i) = d(i)$ and $W = 2m$.

Any reversible Markov chain on the set V can be represented by a (general) random walk on G.



$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ \frac{2}{5} & \frac{2}{5} & 0 & \frac{1}{5} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

$$\boldsymbol{\pi} = \left(\frac{4}{14}, \frac{3}{14}, \frac{5}{14}, \frac{2}{14} \right)$$

Connection with electrical networks

We may build an electrical network with diagram G, in which the edge e has conductance w_e (or, equivalently, resistance $1/w_e$).

Let $s, t \in V$ be distinct vertices termed sources.

Connect a battery across the pair s, t. A current flows along the "wires" of the network. The flow is described by Kirchhoff laws.

We can use electrical networks to analyze any reversible Markov chain.

Infinite Markov chains: recurrence and resistance

Let G be an infinite connected graph with finite vertex-degrees and conductances given by w.

Consider a random walk on G started from a distinguished vertex, say vertex 0.

We want to compute the probability p that the random walk will eventually return to 0.

If $p = 1$, we say that vertex 0 is recurrent. Otherwise, if $p < 1$, it is transient.

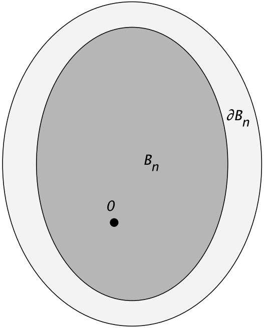
Recurrence and resistance

Let

$$B_n = \{i \in V : d(0, i) \leq n\}$$

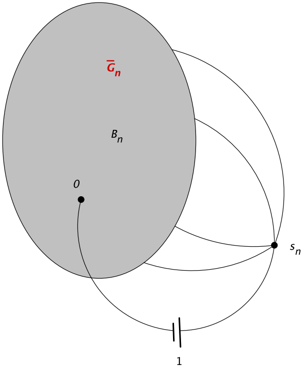
and

$$\partial B_n = B_n \setminus B_{n-1} = \{i \in V : d(0, i) = n\}$$



Consider $G_n = G[B_n]$ and let $\overline{G}_n$ be the graph obtained from G_n by identifying all vertices in ∂B_n in a single vertex s_n.

Regard $\overline{G}_n$ as an electrical network with sources s_n and 0, and let $R_{s0}(n)$ be the effective resistance between these two nodes.



Recurrence and resistance

Observe that $\overline{G}_n$ may be obtained from $\overline{G}_{n+1}$ by identifying all vertices lying in ∂B_n and s_{n+1}.

The identification of two vertices of a network amounts to the addition of a resistor with 0 resistance.

Thus, it can be proved that (Rayleigh principle)

$$R_{s0}(n+1) \geq R_{s0}(n)$$

Therefore, the limit

$$R_{\text{eff}} = \lim_{n\to\infty} R_{s0}(n)$$

exists.

Recurrence and resistance

Let

$$
\begin{aligned}
\phi_n(i) &= \text{P(the randow walk on } G_n \text{ hits } \partial B_n \text{ before } 0 \mid X_0 = 0) \\
&= \text{P(the randow walk on } \overline{G}_n \text{ hits } s_n \text{ before } 0 \mid X_0 = 0)
\end{aligned}
$$

ϕ_n is is the unique harmonic function on G_n satisfying

$$\phi_n(0) = 0, \quad \phi_n(j) = 1 \ \text{ for } j \in \partial B_n$$

ϕ_n is also the potential function on $\overline{G}_n$ viewed as an electrical network with sources $\{s_n, 0\}$.

Recurrence and resistance

Thus,

$$\mathrm{P}(\text{return to 0 before reaching } \partial B_n \mid X_0 = 0)$$

$$= 1 - \sum_{i \in \Gamma(0)} p_{0i}\, \phi_n(i) = 1 - \frac{1}{w(0)} \sum_{i \in \Gamma(0)} w_{0i}\, \phi_n(i)$$

$$= 1 - \frac{1}{w(0)} \text{(current leaving the network at 0)}$$

$$= 1 - \frac{1}{w(0) R_{s0}(n)}$$

Recurrence and resistance

Hence, as $n \to \infty$

$$\mathrm{P}(\text{return to 0 before reaching } \partial B_n \mid X_0 = 0) \to 1 - \frac{1}{w(0)R_{\text{eff}}}$$

On the other hand, by the continuity of probability measures.

$$\mathrm{P}(\text{return to 0 before reaching } \partial B_n \mid X_0 = 0) \\ \to \mathrm{P}(\text{ultimate retun to 0} \mid X_0 = 0)$$

Recurrence and resistance

Theorem

The probability of ultimate return to 0 *is*

$$P(X_n = 0 \text{ for some } n \geq 1 \mid X_0 = 0) = 1 - \frac{1}{w(0)R_{eff}}$$

As in an irreducible Markov chain all the states are either recurrent or transient, the following result holds

Corollary

The random walk is recurrent if and only if $R_{eff} = \infty$.

Random walks in algorithm design

The application of random walks in algorithm design makes use of the fact that (for connected, non-bipartite graphs) the distribution of X_k tends to the stationary distribution $\boldsymbol{\pi}$ as $k \to \infty$.

Moreover, when G is regular $\boldsymbol{\pi}$ is the uniform distribution.

After sufficiently many steps, a node of a random walk in a regular graph is essentially uniformly distributed.

Application to sampling from large sets with complicated structure.

Metropolis algorithm

Metropolis, Rosenbluth, Rosenbluth, Teller and Teller (1953) proposed a simple way to modify the random walk, so that it converges to an arbitrary prescribed probability distribution.

Let G be a regular graph and let $F : V \to \mathbb{R}^+$.

Suppose that at time k we are visiting vertex i.

Choose a random $j \in \Gamma(i)$.

> *If $F(j) \geq F(i)$ then we move to j.*
>
> *Else flip a biased coin and move to j only with probability $F(j)/F(i)$ (and stay at i with probability $1 - F(j)/F(i)$).*

Metropolis algorithm

This modified random walk is again a reversible Markov chain. (It can be considered as a random walk in a graph with edge-weights)

Theorem

The stationary distribution $\boldsymbol{\pi}_F$ of the random walk on G filtered by a function F is given by

$$\boldsymbol{\pi}_F(i) = \frac{F(i)}{\sum_{j \in V} F(j)}, \quad i \in V$$

Random spanning trees

Aldous (1990) and Broader (1989) proposed a very elegant method to select at random (with uniform distribution) a spanning tree on a given graph G.

Start a random walk on G from an initial vertex i.

For each $j \in V$, $j \neq i$, mark the edge through which j is first entered.

Let $\mathcal{T}$ be the set of marked edges.

Theorem

With probability 1, $\mathcal{T}$ is a spanning tree, and every spanning tree occurs with the same probability.