Lecture 7: “Random Walks - II ”

Sotiris Nikoletseas
Associate Professor

CEID - ETY Course
2013 - 2014
Overview

A. Markov Chains

B. Random Walks on Graphs
A. Markov Chains - Stochastic Processes

- **Stochastic Process**: A set of random variables \( \{X_t, t \in T\} \) defined on a set \( D \), where:
  - \( T \): a set of indices representing time
  - \( X_t \): the state of the process at time \( t \)
  - \( D \): the set of states

- The process is discrete/continuous when \( D \) is discrete/continuous. It is a discrete/continuous time process depending on whether \( T \) is discrete or continuous.

- In other words, a stochastic process abstracts a random phenomenon (or experiment) evolving with time, such as:
  - the number of certain events that have occurred (discrete)
  - the temperature in some place (continuous)
Let $S$ a state space (finite or countable). A Markov Chain (MC) is at any given time at one of the states. Say it is currently at state $i$; with probability $P_{ij}$ it moves to the state $j$. So:

$$0 \leq P_{ij} \leq 1 \text{ and } \sum_j P_{ij} = 1$$

The matrix $P = \{P_{ij}\}_{ij}$ is the transition probabilities matrix.

The MC starts at an initial state $X_0$, and at each point in time it moves to a new state (including the current one) according to the transition matrix $P$. The resulting sequence of states $\{X_t\}$ is called the history of the MC.
The memorylessness property

 Clearly, the MC is a stochastic process, i.e. a random process in time.

 the defining property of a MC is its memorylessness, i.e. the random process “forgets” its past (or “history”), while its “future” (next state) only depends on the “present” (its current state). Formally:
\[
\Pr\{X_{t+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_{t-1} = i_{t-1}, X_t = i\} = \Pr\{X_{t+1} = j \mid X_t = i\} = P_{ij}
\]

 The initial state of the MC can be arbitrary.
For states $i, j \in S$, the \underline{\text{t-step transition probability}} from $i$ to $j$ is:

$$P_{ij}^{(t)} = \Pr\{X_t = j | X_0 = i\}$$

i.e. we compute the $(i, j)$-entry of the $t$-th power of transition matrix $P$.

\underline{Chapman - Kolmogorov equations:}

$$P_{ij}^{(t)} = \sum_{i_1, i_2, \ldots, i_{t-1} \in S} \Pr\{X_t = j, \bigcap_{k=1}^{t-1} X_k = i_k | X_0 = i\}$$

$$= \sum_{i_1, i_2, \ldots, i_{t-1} \in S} P_{ii_1} P_{i_1 i_2} \cdots P_{i_{t-1} j}$$
The probability of \textbf{first visit} at state $j$ after $t$ steps, starting from state $i$, is:

$$r_{ij}^{(t)} = \Pr\{X_t = j, X_1 \neq j, X_2 \neq j, \ldots, X_{t-1} \neq j | X_0 = i\}$$

The expected number of steps to arrive for the first time at state $j$ starting from $i$ is:

$$h_{ij} = \sum_{t > 0} t \cdot r_{ij}^{(t)}$$
Visits/State categories

- The probability of a visit (not necessarily for the first time) at state $j$, starting from state $i$, is:
  \[ f_{ij} = \sum_{t>0} r_{ij}^{(t)} \]

- Clearly, if $f_{ij} < 1$ then there is a positive probability that the MC never arrives at state $j$, so in this case $h_{ij} = \infty$.

- A state $i$ for which $f_{ii} < 1$ (i.e. the chain has positive probability of never visiting state $i$ again) is a transient state. If $f_{ii} = 1$ then the state is persistent (also called recurrent).

- If state $i$ is persistent but $h_{ii} = \infty$ it is null persistent. If it is persistent and $h_{ii} \neq \infty$ it is non null persistent.

Note. In finite Markov Chains, there are no null persistent states.
Example (I)

- A Markov Chain

- The transition matrix $P$:

$$P = \begin{bmatrix}
\frac{1}{3} & \frac{2}{3} & 0 & 0 \\
\frac{1}{2} & \frac{1}{3} & \frac{1}{8} & \frac{1}{8} \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

- The probability of starting from $v_1$, moving to $v_2$, staying there for 1 time step and then moving back to $v_1$ is:

$$\Pr\{X_3 = v_1, X_2 = v_2, X_1 = v_2|X_0 = v_1\} =$$

$$= P_{v_1v_2} P_{v_2v_2} P_{v_2v_1} = \frac{2}{3} \cdot \frac{1}{8} \cdot \frac{1}{2} = \frac{1}{24}$$
Example (II)

- The probability of moving from $v_1$ to $v_1$ in 2 steps is:
  \[ P_{v_1v_1}^{(2)} = P_{v_1v_1} \cdot P_{v_1v_1} + P_{v_1v_2} \cdot P_{v_2v_1} = \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{2} = \frac{4}{9} \]
  Alternatively, we calculate $P^2$ and get the (1,1) entry.

- The first visit probability from $v_1$ to $v_2$ in 2 steps is:
  \[ r_{v_1v_2}^{(2)} = P_{v_1v_1} P_{v_1v_2} = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} \]
  while \[ r_{v_1v_2}^{(7)} = (P_{v_1v_1})^6 P_{v_1v_2} = \left( \frac{1}{3} \right)^6 \cdot \frac{2}{3} = \frac{2}{3^7} \]
  and \[ r_{v_2v_1}^{(t)} = (P_{v_2v_2})^{t-1} P_{v_2v_1} = \left( \frac{1}{8} \right)^{t-1} \cdot \frac{1}{2} = \frac{1}{2^{3t-2}} \]
  for $t \geq 1$ (since $r_{v_2v_1}^{(0)} = 0$)
Example (III)

- The probability of (eventually) visiting state $v_1$ starting from $v_2$ is:

$$f_{v_2v_1} = \sum_{t \geq 1} \frac{1}{2^{3t-2}} = \frac{4}{7}$$

- The expected number of steps to move from $v_1$ to $v_2$ is:

$$h_{v_1v_2} = \sum_{t \geq 1} t \cdot r_{v_1v_2}^{(t)} = \sum_{t \geq 1} t \cdot (P_{v_1v_1})^{(t-1)} P_{v_1v_2} = \frac{3}{2}$$

(actually, we have the mean of a geometric distribution with parameter $\frac{2}{3}$)
Irreducibility

- Note: A MC can naturally be represented via a directed, weighted graph whose vertices correspond to states and the transition probability $P_{ij}$ is the weight assigned to the edge $(i, j)$. We include only edges $(i, j)$ with $P_{ij} > 0$.

- A state $u$ is reachable from a state $v$ (we write $v \rightarrow u$) iff there is a path $\mathcal{P}$ of states from $v$ to $u$ with $\Pr\{\mathcal{P}\} > 0$.

- A state $u$ communicates with state $v$ iff $u \rightarrow v$ and $v \rightarrow u$ (we write $u \leftrightarrow v$)

- A MC is called irreducible iff every state can be reached from any other state (equivalently, the directed graph of the MC is strongly connected).
Irreducibility (II)

- In our example, \( v_1 \) can be reached only from \( v_2 \) (and the directed graph is not strongly connected) so the MC is not irreducible.

- Note: In a finite MC, either all states are transient or all states are (non null) persistent.

- Note: In a finite MC which is irreducible, all states are persistent.
Absorbing states

- Another type of states: A state $i$ is absorbing iff $P_{ii} = 1$
  (e.g. in our example, the states $v_3$ and $v_4$ are absorbing)

- Another example:

  ![](image)

  The states $v_0, v_n$ are absorbing
Definition. Let $q^{(t)} = (q_1^{(t)}, q_2^{(t)}, \ldots, q_n^{(t)})$ be the row vector whose $i$-th component $q_i^{(t)}$ is the probability that the MC is in state $i$ at time $t$. We call this vector the state probability vector (alternatively, we call it the distribution of the MC at time $t$).

Main property. Clearly

$$q^{(t)} = q^{(t-1)} \cdot P = q^{(0)} \cdot P^t$$

where $P$ is the transition probability matrix.

Importance: rather than focusing on the probabilities of transitions between the states, this vector focuses on the probability of being in a state.
Definition. A state $i$ called periodic iff the largest integer $T$ satisfying the property

$$q_i(t) > 0 \Rightarrow t \in \{a + kT | k \geq 0\}$$

is largest than 1 ($a > 0$ a positive integer); otherwise it is called aperiodic. We call $T$ the periodicity of the state.

In other words, the MC visits a periodic state only at times which are terms of an arithmetic progress of rate $T$. 
Example: a random walk on a bipartite graph clearly represents a MC with all states having periodicity 2. Actually, a random walk on a graph is aperiodic iff the graph is not bipartite.

Definition: We call aperiodic a MC whose states are all aperiodic. Equivalently, the chain is aperiodic iff (gcd: greatest common divisor):

\[ \forall x, y : \gcd\{ t : P_{xy}^{(t)} > 0 \} = 1 \]
Ergodicity

- Note: the existence of periodic states introduces significant complications since the MC “oscillates” and does not “converge”. The state of the chain at any time clearly depends on the initial state; it belongs to the same “part” of the graph at even times and the other part at odd times.

- Similar complications arise from null persistent states.

- Definition. A state which is non null persistent and aperiodic is called ergodic. A MC whose states are all ergodic is called ergodic.

- Note: As we have seen, a finite, irreducible MC has only non-null persistent states.
Stationarity

- **Definition**: A state probability vector (or distribution) \( \pi \) for which
  \[
  \pi(t) = \pi(t) \cdot P
  \]
is called **stationary distribution**

- Clearly, for the stationary distribution we have
  \[
  \pi(t) = \pi(t+1)
  \]

In other words, when a chain arrives at a stationary distribution it “stays” at that distribution for ever, so this the “final” behaviour of the chain (i.e. the probability of being at any vertex tends to a well-defined limit, independent of the initial vertex). This is why we also call it **equilibrium distribution** or **steady state distribution**. We also say that the chain **converges to stationarity**.
The Fundamental Theorem of Markov Chains

- In general, a stationary distribution may not exist so we focus on Markov Chains with stationarity.

- Theorem. For every irreducible, finite, aperiodic MC it is:
  1. The MC is ergodic.
  2. There is a unique stationary distribution \( \pi \), with \( \pi_i > 0 \) for all states \( i \in S \).
  3. For all states \( i \in S \), it is \( f_{ii} = 1 \) and \( h_{ii} = \frac{1}{\pi_i} \).
  4. Let \( N(i, t) \) the number of times the MC visits state \( i \) in \( t \) steps. Then

\[
\lim_{t \to \infty} \frac{N(i, t)}{t} = \pi_i
\]

Namely, independently of the starting distribution, the MC converges to the stationary distribution.
Definition: A $n \times n$ matrix $M$ is stochastic if all its entries are non-negative and for each row $i$, it is:

$$\sum_j M_{ij} = 1$$

(i.e. the entries of any row add to 1). If in addition the entries of any column add to 1, i.e. for all $j$ it is:

$$\sum_i M_{ij} = 1$$

then the matrix is called doubly stochastic.

Lemma: The stationary distribution of a Markov Chain whose transition probability matrix $P$ is doubly stochastic is the uniform distribution.

Proof: The distribution $\pi_v = \frac{1}{n}$ for all $v$ is stationary, since it satisfies:

$$[\pi \cdot P]_v = \sum_u \pi_u P_{uv} = \sum_u \frac{1}{n} P_{uv} = \frac{1}{n} \sum_u P_{uv} = \frac{1}{n} \sum_u 1 = \pi_v$$
Stationarity in symmetric chains

- **Definition:** A chain is called **symmetric** iff:
  \[ \forall u, v : P_{uv} = P_{vu} \]

- **Lemma:** If a chain is symmetric its stationary distribution is **uniform**.

- **Proof:** Let \( N \) be the number of states. From Fundamental Theorem, it suffices to check that \( \pi_u = \frac{1}{N}, \forall u \), satisfies \( \pi \cdot P = \pi \). Indeed:
  \[
  (\pi P)_u = \sum_v \pi_v \cdot P_{vu} = \frac{1}{N} \sum_v P_{uv} = \frac{1}{N} \cdot 1 = \pi_u \quad \square
  \]
Given a set of $n$ cards, let a Markov Chain whose states are all possible permutations of the cards ($n!$) and one step transition between states defined by some card shuffling rule. For the shuffling to be effective the stationarity distribution must be the uniform one. We provide two such effective shufflings:

(1) Random transpositions: “choose” any two cards at random and swap them e.g.

$$
\cdots a \cdots b \cdots \Rightarrow \cdots b \cdots a \cdots
$$

Note: Indeed the transition probabilities in both directions are the same $\left(\frac{1}{\binom{n}{2}}\right)$ so the chain is symmetric and its stationary distribution uniform.
(2) **Top-in-at-Random**: “place the top card to a random new position of the $n$ possible ones”

**Note**: There are $n$ potential new states. Also, each state can be reached from $n$ other states with probability $\frac{1}{n}$ from each. So the chain is doubly stochastic and its stationary distribution uniform.
On the mixing time

- Although the Fundamental Theorem guarantees that an aperiodic, irreducible finite chain converges to a stationary distribution, it does not tell us how fast convergence happens.

- The convergence rate appropriately close to stationarity is captured by an important measure (the “mixing time”).
As an example, the number of shufflings needed by “Top-in-at-Random” to produce an almost uniform permutation of cards is $O(n \log n)$. Other methods are faster e.g. their mixing time is $O(\log n)$, such as in Riffle-Shuffle where the deck of cards is randomly split into two sets (left, right) which are then “interleaved”.

This convergence rate is very important in algorithmic applications, where we want to ensure that a proper sample can be obtained in fairly small time, even when the state space is very large!
B. Random walks on graphs

- Let $G = (V, E)$ a connected, non-bipartite, undirected graph with $n$ vertices. We define a Markov Chain $MC_G$ corresponding to a random walk on the vertices of $G$, with transition probability:

$$P_{uv} = \begin{cases} \frac{1}{d(u)}, & \text{if } uv \in E(G) \\ 0, & \text{otherwise} \end{cases}$$

where $d(u)$ is the degree of vertex $u$.

- Since the graph is connected and undirected, $MC_G$ is clearly irreducible. Also, since the graph is non-bipartite, $MC_G$ is aperiodic.
The stationary distribution

- So (from fundamental theorem of Markov Chains) $M_G$ has a unique stationary distribution $\pi$.

**Lemma 1**: For all vertices $v \in V$ it is $\pi_v = \frac{d(v)}{2m}$, where $m$ is the number of edges of $G$.

**Proof**: From the definition of stationarity, it must be:

$$\pi_v = [\pi \cdot P]_v = \sum_u \pi_u P_{uv}, \forall v \in V$$

Because of uniqueness, it suffices to verify the claimed solution. Indeed, for all $v \in V$ we have (for the claimed solution value):

$$\sum_u \pi_u P_{uv} = \sum_{u:uv \in E} \frac{d(u)}{2m} \frac{1}{d(u)} = \frac{1}{2m} \sum_{u:uv \in E} 1 = \frac{1}{2m} d(v) = \pi_v$$

$\square$
Definition: The hitting time $h_{uv}$ is the expected number of steps for random walk starting at vertex $u$ to first reach vertex $v$.

Lemma 2: For all vertices $v \in V$, $h_{vv} = \frac{2m}{d(v)}$

Proof: From fundamental theorem:

$$h_{vv} = \frac{1}{\pi_v} = \frac{2m}{d(v)} \quad \text{(from Lemma 1)}$$

Definition: The commute time between $u$ and $v$ is

$$CT_{uv} = h_{uv} + h_{vu}$$

Definition: Let $C_u(G)$ the expected time the walk, starting from $u$, needs to visit every vertex in $G$ at least once. The cover time of the graph, denoted by $C(G)$, is:

$$C(G) = \max_u C_u(G)$$
Lemma: For any edge \((u, v) \in E\) : \(h_{uv} + h_{vu} \leq 2m\)

Proof: Consider a new Markov Chain with states the edges of the graph (every edge taken twice as two directed edges), where the transitions occur between adjacent edges. The number of states is clearly \(2m\) and the current state is the last (directed) edge visited. The transition matrix is

\[
Q(u,v)(v,w) = \frac{1}{d(v)}
\]

This matrix is clearly doubly stochastic since not only the rows but also the columns add to 1. Indeed:

\[
\sum_{x \in V, y \in \Gamma(x)} Q(x,y)(v,w) = \sum_{u \in \Gamma(v)} Q(u,v)(v,w) = \sum_{u \in \Gamma(v)} \frac{1}{d(v)} = d(v) \frac{1}{d(v)}
\]
So the stationary distribution is uniform. So if \( e = (u, v) \) any edge, then \( \pi_e = \frac{1}{2m} \) and \( h_{ee} = \frac{1}{\pi_e} = 2m \). In other words, the expected time between successive traversals of edge \( e \) is \( 2m \).

Consider now \( h_{uv} + h_{vu} \). This is the expected time to go from \( u \) to \( v \) and then return back to \( u \). Conditioning on the event that we initially arrived to \( u \) from \( v \), then \( Q_{(v,u)(v,u)} \) is the time between two successive passages over the edge \( vu \) and is an upper bound to the time to go from \( u \) to \( v \) and back.

But this time is at most \( 2m \) in expectation. Since the MC is memoryless, we can remove the arrival conditioning and the result holds independently of the vertex we initially arrive to \( u \) from.
A resistive electrical network can be seen as an undirected graph. Each edge of the graph is associated to a branch resistance. The electrical flow in the network is governed by two laws:

- Kirchoff’s law for preservation of flow (e.g. all flow that enters a node, leaves it).
- Ohm’s law: the voltage across a resistor equals the product of the resistance times the current through it).

The effective resistance $R_{uv}$ between nodes $u$ and $v$ is the voltage difference between $u$ and $v$ when current of one ampere is injected into $u$ and removed from $v$ (or injected at $v$ and removed from $u$).(Thus, the effective resistance is upper bound by the branch resistance but it can be much smaller).

Given an undirected graph $G$, let $N(G)$ the electrical network defined over $G$, associating 1 Ohm resistance to each of the edges.
**Lemma**: For any two vertices $u, v$ in $G$, the commute time between them is: $CT_{uv} = 2m \cdot R_{uv}$, where $m$ is the number of edges of the graph and $R_{uv}$ the effective resistance between $u$ and $v$ in the associated electrical network $N(G)$.

**Proof**: Let $\Phi_{uv}$ the voltage at $u$ in $N(G)$ with respect to $v$, where $d(x)$ amperes (degree of $x$) of current are injected to each node $x \in V$ and all $2m = \sum_x d(x)$ amperes are removed from $v$. It is:

$$h_{uv} = \Phi_{uv} \quad (1)$$
Indeed, the voltage difference on the edge \( uw \) is
\[
\Phi_{uw} = \Phi_{uv} - \Phi_{wv}.
\]
Using the two laws we get, for all \( u \in V - \{u\} \) that:
\[
d(u) \overset{K}{=} \sum_{w \in \Gamma(u)} \text{current}(uw) \overset{O}{=} \sum_{w \in \Gamma(u)} \frac{\Phi_{uw}}{\text{resistance}(uw)}
\]
\[
= \sum_{w \in \Gamma(u)} (\Phi_{uv} - \Phi_{wv}) = d(u) \cdot \Phi_{uv} - \sum_{w \in \Gamma(u)} \Phi_{wv}
\]
\[
\Rightarrow \Phi_{uv} = 1 + \frac{1}{d(u)} \sum_{w \in \Gamma(u)} \Phi_{wv} \tag{2}
\]
On the other hand, from the definition of expectation we have, for all \( u \in V - \{v\} \), that:

\[
h_{uv} = 1 + \frac{1}{d(u)} \sum_{w \in \Gamma(u)} h_{wv} \tag{3}
\]

Equations (2) and (3) are actually linear systems, with unique solutions (system (2) refers to voltage differences, which are uniquely determined by the current flows). Furthermore, if we identify \( \Phi_{uv} \) in (2) with \( h_{uv} \) in (3), the two systems are identical. This proves that \( h_{uv} = \Phi_{uv} \) indeed (as in (1)).

Now note that \( h_{uv} \) is the voltage \( \Phi_{uv} \) at \( v \) in \( N(G) \) measured w.r.t. \( u \), when currents are injected into all nodes and removed from all other nodes.
Proof (continued)

- Let us now consider a Scenario B, which is like Scenario A except that we remove the $2m$ current units from node $u$ instead of node $v$.

- Denoting the voltage differences in Scenario B by $\Phi'$, we have (as in (1)) that
  \[ \Phi'_{vu} = h_{vu} \]

- Now let us consider a Scenario C, which is like B but with all currents reversed. Denoting the voltage differences in this scenario by $\Phi''$, we have:
  \[ \Phi''_{uv} = -\Phi'_{uv} = \Phi'_{vu} = h_{vu} \]
Finally, consider a Scenario D, which is just the sum of Scenarios A and C. Denoting $\Phi'''$ the voltage differences in D and since the currents (except the $2m$ ones at $u, v$) cancel out, we have

$$\Phi'''_{uv} = \Phi_{uv} + \Phi''_{uv} = h_{uv} + h_{vu}$$

But in D, $\Phi'''_{uv}$ is the voltage difference between $u$ and $v$ when pushing $2m$ amperes at $u$ and removing them at $v$, so (by definition of the effective resistance and Ohm’s law) we have

$$\Phi'''_{uv} = 2m \cdot R_{uv}$$
The line graph. Consider $n + 1$ points on a line:

By symmetry, it is $h_{0n} = h_{n0}$. Also (since the effective resistance between 0 and $n$ is clearly $n$), we have:

$$h_{0n} + h_{n0} = C_{0n} = 2m \cdot R_{0n} = 2 \cdot n \cdot n = 2n^2,$$

thus

$$h_{0n} = h_{n0} = n^2$$

We see that in this case the hitting times are symmetric. This is not the case in general.
Examples (II)

The lollipop graph, composed of a line of $\frac{n}{2} + 1$ vertices joined to a $K_{\frac{n}{2}}$ clique, as in the following figure:

Let $u$ and $v$ the endpoints of the line. We have:

$$h_{uv} + h_{vu} = C_{uv} = 2 \cdot mR_{uv} = 2\Theta(n^2) \cdot \Theta(n) = \Theta(n^3)$$

But from line example in the previous slide

$$h_{uv} = \Theta(n^2) \text{ thus } h_{vu} = \Theta(n^3)$$

This asymmetry is due to the fact that, when we start from $u$, the walk has no option but to go towards $v$; but when we start from $v$ there is very little probability, i.e. $\Theta\left(\frac{1}{n}\right)$, of proceeding to the line.
The cover time

We will now give bounds on the cover time. The first one is rather loose since it is independent of the structure of the graph and only takes into account the number of edges:

**Theorem.** For any connected graph $G(V, E)$, the cover time is:

$$C(G) \leq 2|E||V| = 2 \cdot m \cdot n$$

**Proof.** Consider any spanning tree $T$ of $G$. For any vertex $u$, it is possible to traverse the entire tree and come back to $u$ covering each edge exactly twice:

![Diagram of a spanning tree with a vertex $u$ traversed](image)
The cover time

Clearly, the cover time from vertex $u$ is upper bounded by the expected time for the walk to visit the vertices of $G$ in this order. Let $u = v_0, v_1, \ldots, v_{2n-2} = u$ denote the visited vertices in such a traversal. Then

$$C(u) \leq \sum_{i=0}^{2n-2} h_{v_i,v_{i+1}} = \sum_{(x,y) \in T} (h_{xy} + h_{yx})$$

By the previous lemma on the commute time, we have

$$C(G) = \max_{u \in V} C(u) \leq \sum_{(x,y) \in T} (h_{xy} + h_{yx}) = 2m \sum_{(x,y) \in T} R_{xy} \leq 2 \cdot m \cdot n$$

since for any two adjacent vertices $x, y$ the effective resistance is at most $R_{xy} \leq 1$.

(alternatively we can use a previous Lemma stating that the commute time along an edge is at most $2m$, and the tree has $n - 1$ edges).
Examples

1. The line graph. It has \( n + 1 \) vertices and \( m = n \) edges so
   \[
   C(G) \leq 2 \cdot n(n + 1) \approx 2n^2
   \]
   Also, we know that \( C(G) \geq H_{0n} = n^2 \), thus the bound is tight (up to constants) in this case.

2. The lollipop graph. We get \( C(G) \leq 2 \cdot \Theta(n^2) \cdot n = \Theta(n^3) \).
   Again \( C(G) \geq H_{vu} = \Theta(n^3) \) so the bound is tight.

3. The complete graph. We set \( C(G) \leq 2 \cdot \Theta(n^2) \cdot n = \Theta(n^3) \).
   But from coupon collectors, the cover time is actually
   \[
   C(h) = (1 + o(1))n \ln n, \text{ thus it is much smaller than the upper bound.}
   \]

Comment: This shows a rather counter-intuitive property of cover times (and hitting times): they are not monotonic w.r.t. adding edges to the graph!
Theorem (proof in the book). Let the resistance of a graph $G$ be $R = \max_{u,v \in V} R_{u,v}$. For a connected graph $G$ its cover time is:

$$m \cdot R \leq C(G) \leq c \cdot m \cdot R \cdot \log n$$

for some constant $c$.

Examples:

a) In the complete graph, the probability of hitting a given vertex $v$, when starting at any vertex $u$, is $\frac{1}{n-1}$ so, $\forall u, v \in V$, $h_{uv} = n - 1$. Also, we have $h_{uv} + h_{vu} = 2mR_{uv} \Rightarrow 2(n - 1) = 2 \frac{n(n-1)}{2} R_{uv} \Rightarrow R_{uv} = \frac{2}{n}$ so we get $C(G) \leq c \frac{n(n-1)}{2} \frac{2}{n} \log n = O(n \log n)$, which is tight up to constants.

b) In the lollipop graph, $R = \Theta(n)$ and $m = \Theta(n^2)$, so the upper bound we get is $C(G) \leq O(n^3 \log n)$ which is worse (by a logarithmic factor) from the looser bound.