Homework 9 Solutions

Note: These solutions are not necessarily model answers. Rather, they are designed to be tutorial in nature, and sometimes contain a little more explanation than an ideal solution. Also, bear in mind that there may be more than one correct solution. The maximum total number of points available is 47.

1. (a) Successively multiplying the row vector (1, 0, 0, 0) by P three times, we get the sequence of vectors 2pts $\frac{1}{10}(0, 3, 1, 6)$, $\frac{1}{100}(58, 16, 22, 4)$ and $\frac{1}{1000}(74, 348, 192, 386)$. Thus the probability of being in state 4 after 3 steps starting in state 1 is 0.386. [Note that there is no need to compute the matrix product P^3 .]

A surprising number of people wrote out all paths of length 3 leading from state 1 to state 4 and summed their probabilities. While this is correct, it shows a lack of understanding of linear algebra and its connection with Markov chains. You should be sure that you understand how to solve this problem using matrix-vector multiplication.

(b) Multiplying the row vector $\frac{1}{4}(1, 1, 1, 1)$ by *P* three times as in part (a), we end up with the vector 2pts $\frac{1}{4000}(1003, 1210, 1009, 778)$. Thus the probability of being in state 4 after 3 steps starting in the uniform distribution is $\frac{778}{4000} = 0.1945$.

The same comment as in part (a) applies to this part.

(c) The stationary distribution is $\pi = (0.2346, 0.3048, 0.2631, 0.1975)$ (to 4 decimal places), by solving 2pts for $\pi \mathbf{P} = \pi$ using a linear algebra package. [Note that π is a *left* (row) eigenvector, not the more commonly used right (column) eigenvector. The corresponding right eigenvector with eigenvalue 1 of *any* stochastic matrix is always the uniform distribution!]

Quite a lot of students just took a large enough power of P so that it "looked as if it had converged". As in parts (a) and (b), this indicates a lack of understanding: the point here is that π is the unique (right) eigenvector of P with eigenvalue 1, and you can find this eigenvector using any linear algebra package.

(d) By numerical computation, we see that the variation distance $||p_1^t - \pi|| := \frac{1}{2} \sum_y |P_{1,y} - \pi_y|$ has the 2*pts* value 0.00135 at t = 20 and the value 0.000989 at t = 21. Therefore the desired solution for a threshold of 0.001 is t = 21.

A surprising number of students got the wrong answer here—for reasons that were not entirely clear. Note that this just involves taking large enough powers of P until the above distance drops below 0.001. This is a simple computation using any linear algebra package that supports matrix multiplication.

(a) Recall that, by definition, a state i is recurrent iff ∑_t r^t_{i,i} = 1, where r^t_{i,i} is the probability that the first 6pts return time from state i back to itself is t. Letting T_i denote the return time from i to itself, we can equivalently say that i is recurrent iff Pr[T_i < ∞] = 1. Accordingly, we define f_i := Pr[T_i < ∞]. Now consider the sum ∑_t P^t_{i,i}. Clearly this sum is equal to E[V_i], where V_i is the total number of return visits back to i, starting from i.

Putting together the previous two points, our task is to show that

$$f_i = 1 \iff \mathrm{E}[V_i] = \infty.$$
 (*)

(By the first paragraph, the LHS is equivalent to *i* being recurrent, and by the second paragraph the RHS is equivalent to $\sum_{t} P_{i,i}^{t} = \infty$.)

Finally, to prove (*), we argue as follows. For $j \ge 1$, let $T^{(j)}$ denote the length of the *j*th return visit from *i* back to *i*. Note that the $T^{(j)}$ are mutually independent, and all have the same distribution as T_i .

Now, by the standard tail bound for nonnegative integer-valued r.v.'s, we have

$$E[V_i] = \sum_{k=1}^{\infty} \Pr[V_i \ge k]$$

=
$$\sum_{k=1}^{\infty} \Pr[T^{(1)}, \dots, T^{(k)} \text{ are all finite}]$$

=
$$\sum_{k=1}^{\infty} f_i^k.$$

Clearly, this last sum is finite iff $f_i < 1$. This proves (*) and hence finishes the problem. Several students did not explicitly note that state *i* is transient iff $f_i < 1$. This observation is important, as it demonstrates your understanding of what it means for a state to be transient vs. recurrent.

(b) Clearly $P_{0,0}^{2t}$ is equal to the probability that exactly t of the first 2t steps of the random walk are to the *3pts* right (respectively, to the left). Since the number of steps to the left has distribution Bin(2t, p), we have

$$P_{0,0}^{2t} = \binom{2t}{t} p^t q^t. \tag{**}$$

Using Stirling's formula $n! \sim \sqrt{2\pi n} (\frac{n}{e})^n$, we can get the asymptotic behavior of the binomial coefficient $\binom{2t}{t} \sim \frac{2^{2t}}{\sqrt{\pi t}}$. Thus from (**) we get $P_{0,0}^{2t} \sim \frac{(4pq)^t}{\sqrt{\pi t}}$, as required.

Several students did not explicitly use Stirling's formula to prove that $\binom{2t}{t} \sim \frac{2^{2t}}{\sqrt{\pi t}}$. Although this fact was mentioned in the hint, you were expected to prove it.

(c) Evidently, due to periodicity, $P_{0,0}^{2t+1} = 0$ for all t. Hence by part (b) we get

$$\sum_{t} P_{0,0}^{t} = \sum_{t} P_{0,0}^{2t} = \sum_{t} \frac{(4pq)^{t}}{\sqrt{\pi t}}.$$

In the symmetric case $p = q = \frac{1}{2}$, the above sum is $\sum_t \frac{1}{\sqrt{\pi t}} = \infty$, which by part (a) implies that state 0 is recurrent. On the other hand, if $p \neq \frac{1}{2}$ then $4pq = \alpha < 1$, so the above sum is $\sum_t \frac{\alpha^t}{\sqrt{\pi t}} < \infty$ and by part (a) state 0 is transient.

A lot of students omitted to note that $P_{0,0}^{2t+1} = 0$, i.e., it is impossible to return to 0 at odd time steps.

Note: A similar argument can be used to show that symmetric random walk is also recurrent in 2 dimensions, and is transient in 3 and higher dimensions. This fact is captured in a famous quote by the Japanese mathematician Shizuo Kakutani: "A drunk man will find his way home, but a drunk bird may get lost forever."

(d) Recall that, by definition, state 0 is null recurrent iff the expected return time $E[T_0]$ is infinite. Note *4pts* that we can write this expectation as

$$E[T_0] = 1 + \frac{1}{2}(H_1 + H_{-1}) = 1 + H_1,$$

where H_i denotes the expected hitting time from state *i* to state 0 and we have used symmetry to deduce that $H_1 = H_{-1}$. Thus state 0 is recurrent iff H_1 is infinite.

We now write down a set of equations for the H_i with $i \ge 1$, namely:

$$H_i = 1 + \frac{1}{2}(H_{i-1} + H_{i+1})$$
 or equivalently $H_{i+1} = 2H_i - H_{i-1} - 2$,

where $H_0 = 0$. It is easy to verify by substitution that this family of equations has the solution $H_i = iH_1 - i(i-1)$ for $i \ge 2$, as suggested in the hint. Now suppose for contradiction that H_1 is finite. This would imply that, for some sufficiently large i, H_i is negative. Since this is impossible, we deduce that H_1 must be infinite and therefore state 0 is null recurrent.

2pts

- 3. (a) It follows from the coupon collecting problem that for all v, $C_v(G) = \Theta(n \log n)$. (We are trying to 2pts hit all the vertices, and at each step, we hit each vertex with equal probability, except for the minor detail that we never hit the same vertex twice in succession. It is easy to check that this small change makes no difference to the asymptotics of coupon collecting.)
 - (b) The lollipop graph has n vertices and $O(n^2)$ edges, so by the general bound proved in class and in 2pts MU, Lemma 7.15, the cover time is bounded above by $O(|V| \cdot |E|) = O(n^3)$.

A lot of students attempted to prove that $C(G) = O(n^3)$ in ways similar to the solutions of parts (c) and (d). In particular, they only upper-bounded $C_v(G)$ for a single starting vertex v. This approach is invalid because C(G) is defined as the maximum $C_v(G)$ over all v.

(c) We know from the analysis of random walk on the line graph (MU, pp. 173–174) that $H_{b,a} = \Theta(n^2)$, 4pts so we have the lower bound $C_b(G) = \Omega(n^2)$. (Note that the existence of the clique is irrelevant until the walk reaches a for the first time.) For the upper bound, the same observation tells us that the expected number of steps to reach a from b (and thus along the way to visit all the vertices on the "tail") is $O(n^2)$; to complete the upper bound, we need to show that the walk, starting at a, will visit all vertices of the clique in expected time $O(n^2)$. Starting at a, with probability $\frac{1}{n/2} = \frac{2}{n}$ the walk moves right onto the tail, in which case the time is still bounded by the overall cover time $C(G) = O(n^3)$. With probability $1 - \frac{2}{n}$ the walk moves left onto the clique. At this point, we could essentially use the $\Theta(n \log n)$ cover time for $K_{n/2}$ from part (a), with the following caveat: every time the walk hits a, there is a small chance that it will make an excursion onto the tail. However, in $O(n \log n)$ steps on the clique the expected number of times the walk hits a is only $O(\log n)$, and whenever this happens, with probability $\frac{2}{n}$ it makes an expected O(n) steps on the tail before returning to a. (This is because the expected return time for any vertex in the tail, assuming the first move is to the right, is $\frac{2}{n}$, the reciprocal of the associated stationary probability.) Putting all this together, we get

$$C_b(G) \le H_{b,a} + \frac{2}{n}C(G) + \left(1 - \frac{2}{n}\right) \left(\Theta(n\log n) + O(\log n) \cdot \frac{2}{n} \cdot O(n)\right) = O(n^2),$$

where we have used the general upper bound $C(G) = O(n^3)$ as stated above. Hence $C_b(G) = \Theta(n^2)$, as required.

Note: Technically, we should use Wald's Equation to justify the above argument. Wald's Equation (MU Section 13.3) says that $E[\sum_{i=1}^{T} X_i] = E[T]\mu$, where the X_i are iid r.v.'s with mean μ and T is a (random) stopping time. We are implicitly using this fact above when we bound the total time spent on excursions from a into the tail. The same applies to our analysis in part (d) below. We did not require you to explicitly appeal to Wald's Equation for this problem.

A lot of students did not consider that when the walk reaches vertex a, there is only a 2/n chance that it then goes into the tail. This doesn't actually affect the final bound since the expected time it spends in the tail is only O(n), but it indicates a lack of understanding.

- (d) As in part (c), we consider the two cases for the first step of the walk starting at a:
 - Conditioned on the walk moving right at a (which occurs with probability ²/_n), then with probability at least 1 ²/_n the walk will reach a again before reaching b. This follows from the Gambler's Ruin Problem (as discussed in lecture; see also MU, Section 7.2.1): this is the probability that the first player wins in a fair gambling situation between two players whose initial capitals are ⁿ/₂ 1 dollars and 1 dollar respectively. The walk will then take another H_{a,b} steps to get to b. Hence, conditioned on moving right at a, the expected number of steps to reach b is at least (1 ²/_n)H_{a,b}.
 - Conditioned on the walk moving left at a (which occurs with probability $1 \frac{2}{n}$), then it will take an expected $\Omega(n)$ steps within the clique before returning to a. This follows from looking at a geometric distribution with mean $\frac{1}{n/2-1}$, which is the probability of reaching a from any vertex on the clique in one step. Hence, conditioned on moving left at a, the expected number of steps to reach b is at least $H_{a,b} + \Omega(n)$.

4pts

Combing the two cases, we obtain

$$H_{a,b} \ge \frac{2}{n} \left(1 - \frac{2}{n} \right) H_{a,b} + \left(1 - \frac{2}{n} \right) \left(H_{a,b} + \Omega(n) \right) = H_{a,b} - \frac{4}{n^2} H_{a,b} + \left(1 - \frac{2}{n} \right) \Omega(n) + \frac{1}{n^2} H_{a,b} + \left(1 - \frac{2}{n} \right) \Omega(n) + \frac{1}{n^2} H_{a,b} + \frac{1}{n^2} H_{a,b}$$

as required. Rearranging the terms yields: $\frac{4}{n^2}H_{a,b} \ge \Omega(n)$ and thus $H_{a,b} = \Omega(n^3)$.

- (e) Clearly, $C(L_n) \ge H_{a,b} = \Omega(n^3)$ (using part (d)). Combined with the upper bound from part (b), this *lpt* gives $C(L_n) = \Theta(n^3)$.
- (f) False. For a counterexample, take G to be the lollipop graph, and G' to be the complete graph on n 1pt vertices. Then $\Theta(n \log n) = C(G') < C(G) = \Theta(n^3)$.
- (g) False. Take G to be the line graph on n vertices, and G' to be the lollipop graph. Then $\Theta(n^3) = lpt C(G') > C(G) = \Theta(n^2)$.
- 4. (a) To see that the process is irreducible, note that from any configuration we can reach a configuration 3pts in which all the k particles sit consecutively on the cycle (by successively moving particles clockwise until they hit the next particle around the cycle). Moreover, we can assume that the particles sit in positions $1, \ldots, k$, by moving them all around the cycle as needed. By a symmetrical argument, we can then reach any other configuration from this specific configuration.

Quite a lot of students failed to specify how to move from any configuration to any other. Just saying "the process is irreducible" is not sufficient here (since it's definitely not obvious!).

- (b) The process is aperiodic because some of the configurations have self-loops (i.e., have a non-zero 2pts probability of not changing in one step). These are any configurations in which two or more particles sit consecutively on the cycle. As we have seen in class, a self-loop even on one state in an irreducible chain is enough to ensure aperiodicity (because, starting in any other state, we can reach the self-loop and traverse it any number of times, then return to our starting state).
- (c) We claim that the stationary distribution (which by parts (a) and (b) and the Fundamental Theorem 4pts (MU, Theorem 7.7) must be unique) is the uniform distribution over all $\binom{n}{k}$ configurations. To verify this, fix any configuration σ . Let g be the number of (maximal) consecutive groups of particles in σ . Then the number of (non-self-loop) transitions out of σ is exactly g (one for each group, corresponding to the "head" particle of the group moving clockwise one step). Similarly, the number of (non-self-loop) transitions *into* σ is also g (corresponding to the "tail" particle of each group having moved clockwise one step to its present position). Since all these transitions have the same probability $\frac{1}{k}$ (and since self-loops contribute the same transition probability into and out of σ), it follows that the sum of transition probabilities into σ is equal to the sum of transition probabilities out of σ , which is exactly 1. This means that the Markov chain is "doubly stochastic", so as we saw in class its stationary distribution must be uniform.