Math 135B, Winter 2023.

Homework 7

Due: Mar. 3, 2023

1. Show that the Markov chain in Problem 1 of Homework 4 has a unique invariant distribution and compute it. If, initially, both balls are in urn 1, approximate the probability that one of the urns is empty at time 1000.

2. Consider the Markov chain in Problem 2 of Homework 4.

(a) Show that it has a unique invariant distribution and compute it. What does the matrix P^n look like for large n?

(b) What proportion of time is the state of this chain strictly smaller than the previous state?

3. Consider the Markov chain in Problem 3 of Homework 4.

(a) Show that it has a unique invariant distribution and compute it.

(b) Compute the long-term proportion of tosses on which Alice uses coin 2.

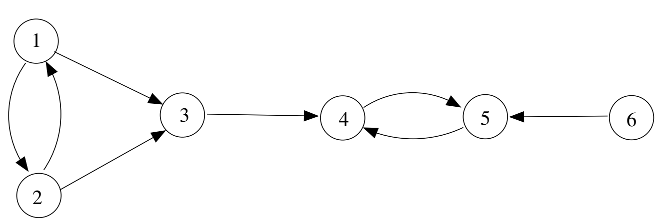
(c) Compute the long-term proportion of tosses on which Alice tosses Heads.

(d) Compute the long-term proportion of tosses on which Alice tosses Heads with coin 2.

(e) Bob pays Alice \$1 every time she tosses Heads with coin 1 and \$2 every time she tosses Heads with coin 2. Determine Alice's long-term average winnings per toss.

4. (A Wall Street job interview question.) You roll a fair six-sided die repeatedly and sum the numbers you roll. What is the expected number of rolls until the sum is a nonzero multiple of 4 for the first time?

5. Currently, it is not public knowledge how Google ranks the web pages in its search algorithm. The initial approach by S. Brin and L. Page in the mid 1990s, known as the *PageRank* algorithm, was using Markov chains for this purpose. Form an oriented graph, with N vertices, which are the web pages, in which an oriented edge goes from a web page x to a web page y if x contains a hyperlink to y. This creates a graph as in the following example with $N = 6$.



A random surfer begins at some page and then at every step chooses one of the outgoing arrows from the current page uniformly at random and then follows it to the next page; if there are no such such arrows, the surfer chooses one of the N pages uniformly at random and moves there. Let W be the transition matrix for the resulting Markov chain.

(a) Write down the matrix W for this example. Determine the classes and their recurrence or transience.

We see that the immediate problem is that the chain may not be irreducible. To deal with this problem, we introduce "damping." Pick a small $\epsilon > 0$. Then, the surfer follows the above algorithm with probability $1 - \epsilon$. Otherwise, with probability ϵ, it chooses a random state, i.e., evolves for one step according to the $N \times N$ transition matrix R that has all entries equal to $1/N$. A typical choice is $\epsilon = 0.15$; assume this for the rest of the problem. This gives the transition matrix

$$P = (1 - \epsilon)W + \epsilon R.$$

Using this transition matrix, the surfer's position has a unique invariant distribution, which is used to rank pages: the more often, in the long run, the surfer visits a web page x, the higher the rank of x.

(b) Rank the pages for the given example using this algorithm.

One problem with this version of *PageRank* is its vulnerability to the *Sybil attacks*. (The name comes from a famous, but possibly fictional, psychiatric case of a woman with multiple personality disorder.) That is, the owner of a page x can improve its rank by creating a number of new web pages $x_1, \ldots, x_k$, with links $x \to x_i$ and $x_i \to x$, for all $i = 1, \ldots, k$. In our example, say, the owner of page 5 creates a single new page 7 (so $k = 1$) with links $5 \to 7$ and $7 \to 5$ (but no other new links).

(c) Using *PageRank*, determine the new ranking of pages after the described Sybil attack in our example.

Math 135B, Winter 2023.

Homework 7 Solutions.

1. Show that the Markov chain in Problem 1 of Homework 4 has a unique invariant distribution and compute it. If, initially, both balls are in urn 1, approximate the probability that one of the urns is empty at time 1000.

Solution. As all entries of P^2 are nonzero, the chain is irreducible and aperiodic. In this case, P is a symmetric matrix, so it is doubly stochastic and $\pi = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$. As $n \to \infty$, $P_{ij}^n \to 1/4$ for all states i and j, by the convergence theorem. Because

$$P(X_n \text{ is in state 1 (11) or 4 (22)} \mid X_0 = 1) = P_{11}^n + P_{14}^n \to 1/4 + 1/4 = 1/2,$$

the probability is approximately 1/2.

2. Consider the Markov chain in Problem 2 of Homework 4.

(a) Show that it has a unique invariant distribution and compute it. What does the matrix P^n look like for large n?

Solution. As all entries of P are nonzero, the chain is irreducible and aperiodic. The invariant distribution is

$$\pi = \begin{bmatrix} 27/89 & 32/89 & 30/89 \end{bmatrix} \approx \begin{bmatrix} 0.3034 & 0.3596 & 0.3371 \end{bmatrix}.$$

The powers P^n converge to 3×3 matrix will all rows equal to π.

(b) What proportion of time is the state of this chain strictly smaller than the previous state?

Solution. The answer is

$$\pi_3(P_{31} + P_{32}) + \pi_2 P_{21} = \frac{30}{89} \cdot \frac{5}{6} + \frac{32}{89} \cdot \frac{1}{4} = \frac{33}{89} \approx 0.3708.$$

3. Consider the Markov chain in Problem 3 of Homework 4.

(a) Show that it has a unique invariant distribution and compute it.

Solution. Irreducibility holds because transitions $1 \to 2 \to 4 \to 3 \to 1$ happen with positive probability. Aperiodicity holds because $1 \to 1$ happens with positive probability. The invariant distribution is

$$\pi = \begin{bmatrix} 49/117 & 28/117 & 28/117 & 4/39 \end{bmatrix} \approx \begin{bmatrix} 0.4188 & 0.2393 & 0.2393 & 0.1026 \end{bmatrix}.$$

(b) Compute the long-term proportion of tosses on which Alice uses coin 2.

Solution. This is the same as proportion of time spent at the state 1 (HH), which is $\pi_1 = 49/117 \approx 0.4188$.

(c) Compute the long-term proportion of tosses on which Alice tosses Heads.

Solution. This is

$$\pi_1 + \pi_3 = 77/117 \approx 0.6581.$$

(d) Compute the long-term proportion of tosses on which Alice tosses Heads with coin 2.

Solution. (Not the product of (b) and (c)!) This is

$$\pi_1 \cdot 0.6 \approx 0.2513.$$

(e) Bob pays Alice \$1 every time she tosses Heads with coin 1 and \$2 every time she tosses Heads with coin 2. Determine Alice's long-term average winnings per toss.

Solution. The answer is

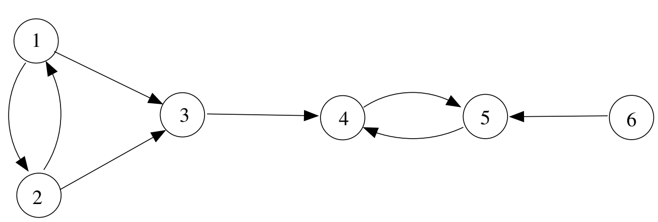
$$2\pi_1 \cdot 0.6 + (1 - \pi_1) \cdot 0.7 = 0.7 + 0.5\pi_1 = 0.9094.$$

4. (A Wall Street job interview question.) You roll a fair six-sided die repeatedly and sum the numbers you roll. What is the expected number of rolls until the sum is a nonzero multiple of 4 for the first time?

Solution. If X_n is your sum modulo 4 after n rolls, then X_n is a Markov chain on the states 0, 1, 2, 3 and the question asks for the expected value of the return time to 0. If we can compute the invariant distribution $\pi = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 \end{bmatrix}$, then the answer will be $1/\pi_0$. Modulo 4, the chain adds to the current state: 0 with probability 1/6 (when you roll 4); 1 with probability 1/3 (when you roll 1 or 5); 2 with probability 1/3 (when you roll 2 or 6); and 3 with probability 1/6 (when you roll 3). The transition matrix can be easily written down (see below), but one could see in advance that it is doubly stochastic, as the matrix R_k of the deterministic chain that adds k is doubly stochastic for any k (as a permutation matrix). So, $\pi = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$, and the answer is therefore 4. To check directly that P is doubly stochastic:

$$P = \frac{1}{6}I + \frac{1}{3}R_1 + \frac{1}{3}R_2 + \frac{1}{6}R_3 = \begin{bmatrix} 1/6 & 1/3 & 1/3 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/3 & 1/3 & 1/6 & 1/6 \end{bmatrix}.$$

5. Currently, it is not public knowledge how Google ranks the web pages in its search algorithm. The initial approach by S. Brin and L. Page in the mid 1990s, known as the *PageRank* algorithm, was using Markov chains for this purpose. Form an oriented graph, with N vertices, which are the web pages, in which an oriented edge goes from a web page x to a web page y if x contains a hyperlink to y. This creates a graph as in the following example with $N = 6$.



A random surfer begins at some page and then at every step chooses one of the outgoing arrows from the current page uniformly at random and then follows it to the next page; if there are no such such arrows, the surfer chooses one of the N pages uniformly at random and moves there. Let W be the transition matrix for the resulting Markov chain.

(a) Write down the matrix W for this example. Determine the classes and their recurrence or transience.

<u>Solution</u>. We have

$$W = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

There are three transient classes: $\{1, 2\}$, $\{3\}$ and $\{6\}$, and one recurrent class $\{4, 5\}$.

We see that the immediate problem is that the chain may not be irreducible. To deal with this problem, we introduce "damping." Pick a small $\epsilon > 0$. Then, the surfer follows the above algorithm with probability $1 - \epsilon$. Otherwise, with probability ϵ, it chooses a random state, i.e., evolves for one step according to the $N \times N$ transition matrix R that has all entries equal to $1/N$. A typical choice is $\epsilon = 0.15$; assume this for the rest of the problem. This gives the transition matrix

$$P = (1 - \epsilon)W + \epsilon R.$$

Using this transition matrix, the surfer's position has a unique invariant distribution, which is used to rank pages: the more often, in the long run, the surfer visits a web page x, the higher the rank of x.

(b) Rank the pages for the given example using this algorithm.

<u>Solution</u>. The left eigenvector of $P = 0.85 \cdot W + 0.15 \cdot R$ is, to 4 decimals:

$$\pi_0 = \begin{bmatrix} 0.0435 & 0.0435 & 0.0620 & 0.4215 & 0.4046 & 0.0250 \end{bmatrix},$$

so that pages are ranked, from first to last: $4, 5, 3, (1, 2), 6$ (with a tie between 1 and 2).

Note. Instead of solving for π_0, the algorithm computes the power P^n for a suitably large n. It can be shown[1] that $|P^n_{ij} - \pi_j| \leq (1 - \epsilon)^n$ for all i, j, n, which, remarkably, is independent of the number of web pages N. Even for billions of pages, taking n to be a few hundred suffices.

One problem with this version of *PageRank* is its vulnerability to the *Sybil attacks*. (The name comes from a famous, but possibly fictional, psychiatric case of a woman with multiple personality disorder.) That is, the owner of a page x can improve its rank by creating a number of new web pages $x_1, \ldots, x_k$, with links $x \to x_i$ and $x_i \to x$, for all $i = 1, \ldots, k$. In our example, say, the owner of page 5 creates a single new page 7 (so $k = 1$) with links $5 \to 7$ and $7 \to 5$ (but no other new links).

(d) Using *PageRank*, determine the new ranking of pages after the described Sybil attack in our example.

[1]See Proposition 10.5 in this book: E. Behrends, "Introduction to Markov Chains With Special Emphasis on Rapid Mixing," Springer, 2000.

Solution. Now,

$$\pi_0 = \begin{bmatrix} 0.0373 & 0.0373 & 0.0531 & 0.2418 & 0.4124 & 0.0214 & 0.1967 \end{bmatrix},$$

so the new rank is: $5, 4, 7, 3, (1, 2), 6$. The owner of the page 5 made sure that it is now well in the first place.