MATH 5040/6810: Homework 1 (Due Mon, Sep. 9)

Problem 1

a) Let $A_0, A_1, A_2, \ldots$ be events such $\mathbb{P}(A_0 \cap \ldots \cap A_n) > 0$ for all integers $n \geq 0$. Show that for all integers $n \geq 1$

$$\mathbb{P}(A_0 \cap \cdots \cap A_n) = \mathbb{P}(A_0)\mathbb{P}(A_1|A_0)\cdots\mathbb{P}(A_n|A_0 \cap \cdots \cap A_{n-1}).$$

Solution: For $n = 1$, the equation says $\mathbb{P}(A_0 \cap A_1) = \mathbb{P}(A_0)\mathbb{P}(A_1|A_0)$, which is true by the definition of conditional probability. Assume now that the equation is true for an arbitrary integer n. We will prove that then it is true for $n + 1$. Indeed,

$$\begin{aligned}
&\mathbb{P}(A_0)\mathbb{P}(A_1|A_0)\cdots\mathbb{P}(A_n|A_0 \cap \cdots \cap A_{n-1})\mathbb{P}(A_{n+1}|A_0 \cap \cdots \cap A_n) \\
&= \mathbb{P}(A_0 \cap \cdots \cap A_n)\mathbb{P}(A_{n+1}|A_0 \cap \cdots \cap A_n) \\
&= \mathbb{P}(A_0 \cap \cdots \cap A_{n+1}).
\end{aligned}$$

Explanation: the first equality comes from the induction assumption that the equation is valid for n and the second equality uses the definition of conditional probability.

Since we proved that the equation is true for $n = 1$ and that if it is true for an integer n, then it is true for the integer $n+1$, we conclude that the equation is true for all integers n. (E.g. if we want to conclude that it is true for $n = 4$, then we argue that since it is true for $n = 1$, it is true for $n = 1 + 1 = 2$, and hence also for $n = 2 + 1 = 3$, and, therefore, for $n = 3 + 1 = 4$ as well.)

b) Consider a stochastic process $\{X_0, X_1, \ldots\}$ on a state space $S = \{a, b, c\}$. What does the equation in part a) say about the probability $\mathbb{P}\{X_0 = a, X_1 = b, X_2 = c, X_3 = b, X_4 = c\}$?

Solution:

$$\begin{aligned}
&\mathbb{P}\{X_0 = a, X_1 = b, X_2 = c, X_3 = b, X_4 = c\} \\
&= \mathbb{P}\{X_0 = a\}\mathbb{P}\{X_1 = b|X_0 = a\}\mathbb{P}\{X_2 = c|X_0 = a, X_1 = b\} \\
&\qquad \times \mathbb{P}\{X_3 = b|X_0 = a, X_1 = b, X_2 = c\}\mathbb{P}\{X_4 = c|X_0 = a, X_1 = b, X_2 = c, X_3 = b\}.
\end{aligned}$$

c) How does the equation you got in part b) simplify if we assume the process in part b) to be a Markov process?

Solution:

$$\begin{aligned}
&\mathbb{P}\{X_0 = a, X_1 = b, X_2 = c, X_3 = b, X_4 = c\} \\
&= \mathbb{P}\{X_0 = a\}\mathbb{P}\{X_1 = b|X_0 = a\}\mathbb{P}\{X_2 = c|X_1 = b\}\mathbb{P}\{X_3 = b|X_2 = c\}\mathbb{P}\{X_4 = c|X_3 = b\}.
\end{aligned}$$

Problem 2

Consider the following Markov chain on state space $\{a, b, c\}$: the chain starts at a, b, or c with probability, respectively, 1/2, 1/6, and 1/3. Then, from a the chain either stays at a with probability 1/3, or goes to c; from b the chain goes to a or c with probability 1/2 each; and from c the chain goes to a, b, or c, with probabilities 1/2, 1/4, and 1/4, respectively.

a) Write the transition matrix of the Markov chain.

Solution:

$$P = \begin{bmatrix} 1/3 & 0 & 2/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{bmatrix}$$

b) Using the formula you developed in 1c) compute the probability $\mathbb{P}\{X_0 = a, X_1 = c, X_2 = b, X_3 = c, X_4 = b\}$.

Solution:

$$\begin{aligned} &\mathbb{P}\{X_0 = a, X_1 = c, X_2 = b, X_3 = c, X_4 = b\} \\ &= \mathbb{P}\{X_0 = a\}\mathbb{P}\{X_1 = c|X_0 = a\}\mathbb{P}\{X_2 = b|X_1 = c\}\mathbb{P}\{X_3 = c|X_2 = b\}P\{X_4 = b|X_2 = c\} \\ &= 1/2 \times 2/3 \times 1/4 \times 1/2 \times 1/4. \end{aligned}$$

c) Compute the probability $\mathbb{P}\{X_1 = c\}$ by considering all the possible scenarios leading to the event $X_1 = c$.

Solution:

$$\begin{aligned} \mathbb{P}\{X_1 = c\} &= \mathbb{P}\{X_0 = a, X_1 = c\} + \mathbb{P}\{X_0 = b, X_1 = c\} + \mathbb{P}\{X_0 = c, X_1 = c\} \\ &= \mathbb{P}\{X_0 = a\}\mathbb{P}\{X_1 = c|X_0 = a\} + \mathbb{P}\{X_0 = b\}\mathbb{P}\{X_1 = c|X_0 = b\} + \mathbb{P}\{X_0 = c\}\mathbb{P}\{X_1 = c|X_0 = c\} \\ &= 1/2 \times 2/3 + 1/6 \times 1/2 + 1/3 \times 1/4. \end{aligned}$$

d) Compute $\mathbb{P}\{X_0 = a|X_1 = c\}$.

Solution:

$$\begin{aligned} \mathbb{P}\{X_0 = a|X_1 = c\} &= \frac{\mathbb{P}\{X_0 = a, X_1 = c\}}{\mathbb{P}\{X_1 = c\}} = \frac{\mathbb{P}\{X_0 = a\}\mathbb{P}\{X_1 = c|X_0 = a\}}{\mathbb{P}\{X_1 = c\}} \\ &= \frac{1/2 \times 2/3}{1/2 \times 2/3 + 1/6 \times 1/2 + 1/3 \times 1/4}. \end{aligned}$$

e) Compute the probability $\mathbb{P}\{X_5 = b\}$. (Hint: do NOT consider all possible scenarios leading to the event $X_5 = b$! It is OK to use a calculator to compute powers of matrices.)

Solution: Using a computer (e.g. matlab or R) we get

$$P^5 = \begin{bmatrix} 0.4285 & 0.1129 & 0.4586 \\ 0.4286 & 0.1145 & 0.4569 \\ 0.4286 & 0.1155 & 0.4559 \end{bmatrix}.$$

The initial condition is $\Phi_0 = [1/2\ 1/6\ 1/3]$. So

$$\Phi_5 = \Phi_0 P^5 = [1/2\ 1/6\ 1/3] \begin{bmatrix} 0.4285 & 0.1129 & 0.4586 \\ 0.4286 & 0.1145 & 0.4569 \\ 0.4286 & 0.1155 & 0.4559 \end{bmatrix} = [0.4286\ 0.1141\ 0.4574].$$

$\mathbb{P}\{X_5 = b\}$ is the second entry of Φ_5 and is, therefore, equal to 0.1141. To put it in other words, recall that the entries of P^5 are the 5-step transition probabilities. So

$$\begin{aligned}
\mathbb{P}\{X_5 = b\} &= \mathbb{P}\{X_0 = a, X_5 = b\} + \mathbb{P}\{X_0 = b, X_5 = b\} + \mathbb{P}\{X_0 = c, X_5 = b\} \\
&= \mathbb{P}\{X_0 = a\}\mathbb{P}\{X_5 = b|X_0 = a\} + \mathbb{P}\{X_0 = b\}\mathbb{P}\{X_5 = b|X_0 = b\} + \mathbb{P}\{X_0 = c\}\mathbb{P}\{X_5 = b|X_0 = c\} \\
&= 1/2 \times 0.1129 + 1/6 \times 0.1145 + 1/3 \times 0.1155,
\end{aligned}$$

which is exactly the second entry of the row vector $\Phi_0 P^5$.

f) Compute the invariant measure of the Markov chain. (OK to use a computer or a calculator.)

Solution: Using a computer (e.g. matlab or R) we can get the right eigenvectors for any given square matrix. But the invariant measure we are after is a left eigenvector of the matrix P. It is thus a right eigenvector of the transpose matrix P'. Asking Matlab for the right eigenvector of the matrix P', corresponding to the eigenvalue 1 gives us [0.6728 0.1794 0.7177]. Normalizing by $0.6728+0.1794+0.7177 = 1.57$ (so that the entries add up to 1) gives the invariant measure $\Phi_\infty = [0.4286\ 0.1143\ 0.4571]$.

g) Compute the limit of $\mathbb{P}\{X_n = b\}$ as $n \to \infty$. (Hint: Invariant measure.)

Solution: The Markov chain is irreducible (because with positive probability we can get from a to c to b to a) and aperiodic (because for example a can go to a). Thus, the Perron-Frobenius theorem tells us that the limit in question is the value of the invariant measure at b. That is 0.1143.

Problem 3

You have a bag with 5 red balls and 3 blue ones. You also have an infinite reserve of red and blue balls. Every minute you draw a ball at random from the bag and look at its color. Then you return it to the bag and add with it, to the bag, another ball of that same color (which you take from the infinite reserve). Let X_n denote the color of the ball you drew at time n. Is X_n a Markov chain? Why or why not? If it is, write its transition matrix.

Solution: This is not a Markov chain. To see that, consider the following two scenarios. In the first scenario, we run the process for $n = 100$ steps and happen to get a red ball each time. Then $X_{100} =$"red" and we have 105 red balls and 3 blue ones. The probability that the next ball is red, given this scenario, is 105/108. In the second scenario, we run the process for $n = 100$ steps and happen to get a blue ball the first 99 times and then a red ball. Then X_{100}="red" and we have 6 red balls and 102 blue balls. The probability that the next ball is red, given this scenario, is $6/108 \neq 105/108$. If this were a Markov chain, then to figure out the probability of X_{101}="red" we would only care about the fact that X_{100}="red", regardless of what happened in the previous 99 steps. But our computation shows that the probabilities are in fact different and do depend on what happened before time 100. So the process is not a Markov chain.

Problem 4

You have \$4 in your pocket. You are given a coin that lands heads with probability p. Every minute you toss the coin once. If it is heads you gain an extra \$1. If it is tails you lose a \$1. If you reach \$0 then you will have \$0 for the rest of the game, regardless of what your coin comes up. The (small) House has \$6. So if you reach \$10 you will have \$10 for the rest of the game regardless of what your coin lands.

Is this a Markov chain? Why or why not? If it is, write its transition matrix.

Solution: Yes this is a Markov chain: if we know the current amount in our pocket, then the probability mass function of the amount after the next move is completely determined. For example, if we are at \$5 we know that at the next step we will have \$6, with probability p, or \$4, with probability $1-p$. The transition matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-p & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-p & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-p & 0 & p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Problem 5 (See the Google Math slides on the course webpage for a similar problem)

Consider a tiny web consisting of 6 sites. Site 1 is a John's homepage. John teaches calculus and his site points to sites 2 and 3. Site 2 is the course syllabus and does not point to any other sites. Site 3 is the calculus course website. It points back at John's homepage and also at the course syllabus. It also points at Emily's webpage, site 5. (Emily is the TA for the course.) Site 4 belongs to a friend of Emily's, Jack. It points at Emily's website and at Jack's old website, site 6. Emily's website points at both Jack's pages, the new one 4 and the old one 6. Jack's old webpage 6 points at his new webpage 4.

a) Draw the graph corresponding to the above network.

b) Consider the stochastic process where from each site you go equally likely to any of the sites it points to. What happens when you eventually reach site 2?

c) Extend the stochastic process by the rule: from a site that does not point at any other site go equally likely to any site on the web (including the original site itself). Does this process give a Markov chain? Write its transition matrix. What are its communication classes? Which are transient and which are recurrent? Compute the invariant measure. Does it give positive probability to all sites? Can we use it to rank the transient states? Why or why not? (It is OK to use a computer or a laptop to compute the probability measure.)

d) To overcome the problem with transient states, consider instead the following Markov chain: from a site that does not point to any other site, go equally likely to any site on the web (including the site itself); but from a site that does points somewhere, flip a coin that gives heads with probability 15%. If the coin lands heads, go equally likely to any site on the web (including the site itself). If, on the other hand, it lands tails then go equally likely to any site that the current site points to. Write the transition matrix for this chain. What are its communication classes? Which are transient and which are recurrent? Compute the invariant measure. Does it give positive probability to all sites?

Solution: The new matrix is

$$0.15 \times \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix} + 0.85 \times \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Its entries are all positive and so this Markov chain is irreducible and aperiodic and its invariant measure must have all positive entries. I used Matlab to compute the left eigenvector corresponding to eigenvalue one. If the software you are using can only compute right eigenvectors, then you can compute the right eigenvector of the transpose of the matrix. This will be the same as the left eigenvector of the matrix. This gave the vector

$$[-0.1044 \;\; -0.1488 \;\; -0.1160 \;\; -0.7043 \;\; -0.4038 \;\; -0.5425].$$

Normalizing by the total sum (to get entries that add up to one) we get the invariant measure

$$[0.0517 \; 0.0737 \; 0.0574 \; 0.3487 \; 0.1999 \; 0.2686].$$

e) Use your results in part d) to rank the six sites.

Solution: Sites with a higher value of the invariant measure are visited more often. Therefore, the ranking, from higher to lower, is: 4, 6, 5, 2, 3, 1.