ISYE 7201: Production & Service Systems
Spring 2022
Instructor: Spyros Reveliotis
1st Midterm Exam (Take Home)
Release Date: February 9, 2022
Due Date: February 13, 2022

While taking this exam, you are expected to observe the Georgia Tech Honor Code. In particular, no collaboration or other interaction among yourselves is allowed while taking the exam.

Please, send me your responses as a pdf file attached to an email. Name the pdf file by your last name (only). The pdf file can be a scan or photos of a hand-written document, but, please, write your answers clearly and thoroughly. Also, make sure that the pdf file is not too big; you can reduce the size of your file by loading it into Adobe Acrobat and saving it with the "reduced" size option before emailing it to me.

Finally, report any external sources (other than your textbook) that you referred to while preparing the solutions.

Problem 1 (20 points): Consider a 3-state discrete-time Markov chain with states 1, 2 and 3. Let X_k denote the process state at period $k = 0, 1, 2, \ldots$. The one-step transition probability matrix for this chain is

$$P = \begin{bmatrix} 0.4 & 0.4 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.1 & 0.5 & 0.4 \end{bmatrix}$$

i. **(5 pts)** Draw the state transition diagram of this Markov chain.

ii. **(5 pts)** Assuming that $X_0 = 2$, predict the process state in periods 1 and 2.

iii. **(5 pts)** Explain that this process has a limiting distribution π, and compute this distribution.

iv. **(5 pts)** If $X_k = 1$, what is the expected number of periods until the process gets into state 3?

Problem 2 (30 points): Consider a rental company that serves (has offices in) two different locations, A and B. A rental contract signed up at location A will return the car at the same location with probability 0.6 and at location B with probability 0.4. A rental contract signed up at location B will return the car at the same location with probability 0.7 and at location A with probability 0.3. Rentals returning the car at the same location generate a profit at a rate of \$15.0 per rental day for the company, and rentals returning the car at the other location generate a profit of \$25.0 per rental day. Cars are not transferred between the two locations by the company itself. What is the average profit rate per rental day for the company?

Problem 3 (20 points) Prove that if two states i and j of a DT-MC communicate, then they have the same period (this is a formal way to establish the (a-)periodicity is a "class property").

Problem 4 (20 points) Consider an $n \times n$ nonnegative matrix A, and also let $G(A)$ be a directed graph with node set $\mathcal{N} = \{1, 2, \ldots, n\}$, and edge set $\mathcal{E} = \{(i, j) \in \mathcal{N} \times \mathcal{N} : a_{i,j} > 0\}$. Also, let $a_{i,j}^{(k)}$ denote the (i, j) element of matrix A^k, $k = 1, 2, \ldots$.

You must prove the following statements:

1. **(10 pts)** For any pair $(i, j) \in \mathcal{N} \times \mathcal{N}$, there is a directed path of length k (i.e., involving k edge-traversals) leading from node i to node j in $G(A)$ if and only if $a_{i,j}^{(k)} > 0$. (Also, assume that a path can revisit some nodes and edges of $G(A)$; more formally, a path that involves such repetitions is known as a *walk* in $G(A)$.)

2. **(10 pts)** If A^k is strictly positive (element-wise) for some $k \geq 1$, then $A^{k'} > 0$ for every $k' \geq k$.

Remark: It is also interesting to notice that the above properties can be established by considering only the $n \times n$ binary matrix $I(A)$, with $I(A)[i, j] \equiv I_{\{a_{i,j} > 0\}}$; i.e., these properties depend only on the *structure* of the considered matrix A with respect to the placement of its non-zero elements, and not on the actual values of these elements.

Problem 5 (20 points): Consider a scoring process involving n experts that iterates as follows: At the first iteration, each expert $i \in \{1, \ldots, n\}$ comes up with a score $s_i(1)$ and broadcasts this score to every other expert. At each subsequent iteration $t \geq 2$, expert i revises her score based on the scores that she has received from the other experts in the previous round, according to an averaging mechanism that is defined as follows:

$$s_i(t) = \sum_{j=1}^{n} w_{i,j} \cdot s_j(t-1)$$

where $w_{i,j} \geq 0$; $w_{i,i} > 0$; and $\sum_{j=1}^{n} w_{i,j} = 1.0$ (i.e., each expert computes her current score as a weighted average of her own score and the scores that were received by the other experts in the previous round). The revised scores are broadcasted again, and the process proceeds to the next iteration.

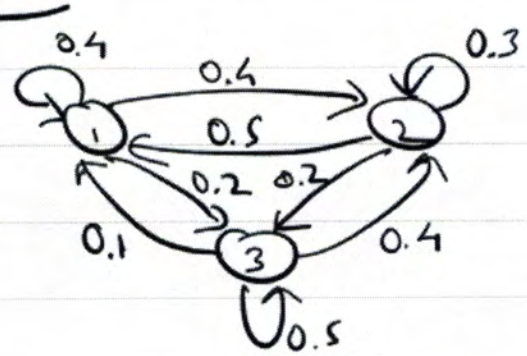
What is the meaning of the weights $w_{i,j}$ in this computation? Your task is to use the results on DT-MCs presented in class in order to identify conditions on this voting scheme so that the experts reach *consensus* (i.e., agree eventually on a common score). Try to come up with as general a condition as possible, and also provide an intuitive explanation of your findings.

Finally, please, explain clearly all your answers.

ISyE 7201: Production & Service Systems
Spring 2022
Midterm I Solutions

Problem 1:

(i)



(ii) $P_0 = (0\ 1\ 0)$

$P_1 = P_0 \cdot P = (0.5\ \ 0.3\ \ 0.2)$

$P_2 = P_1 \cdot P = (0.37\ \ 0.39\ \ 0.24)$

(iii) It is clear from the state transition diagram (STD) that the process is irreducible, and since it has a finite state space, it is positive recurrent. Furthermore, the presence of self-loops in the STD implies the aperiodicity of the process. Hence, the process is ergodic and has a limiting distribution π.

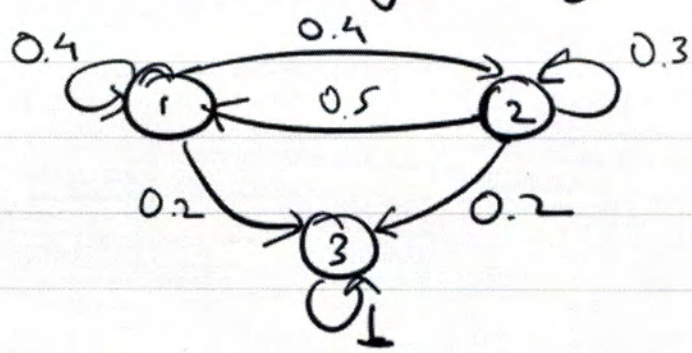
π is provided by the following system of equations

$$\begin{cases} \pi = \pi \cdot P \\ \pi \cdot \underline{1} = 1 \end{cases}$$

The solution of this system of equations is

$$\pi_1 = \frac{16}{44}, \quad \pi_2 = \frac{17}{44}, \quad \pi_3 = \frac{11}{44}$$

(iv) In order to answer this question we need to turn state 3 into an absorbing state, by considering the following DT-MC:



Then, we can see that at each transition from states 1 and 2, the process can end up at state 3 w.p. 0.2; i.e., each of these transitions can be perceived as a Bernoulli trial with success prob. 0.2. Hence, the expected # of periods to get to state 3 is $1/0.2 = 5$.

A more routine approach to answer this question is as follows:

Consider the modified DT-MC, and let:

- $y_1 = E[\#$ of periods for absorption to state $3 \mid X_0 = 1]$
- $y_2 = E[$ " " " " " " " " $\mid X_0 = 2]$

Then, by conditioning upon the first transition from state i, $i = 1, 2$, we have:

$$y_1 = 1 \cdot 0.2 + (1 + y_1)\, 0.4 + (1 + y_2) \cdot 0.4 \Rightarrow$$

$$\Rightarrow 0.6\, y_1 - 0.4 \cdot y_2 = 1 \qquad (1)$$

$$y_2 = 1 \cdot 0.2 + (1 + y_1)\, 0.5 + (1 + y_2) \cdot 0.3 \Rightarrow$$

$$\Rightarrow -0.5\, y_1 + 0.7\, y_2 = 1 \qquad (2)$$

From the linear system of equations defined by (1) and (2), using Crammer's rule, we have:

$$y_1 = \frac{\begin{vmatrix} 1 & -0.4 \\ 1 & 0.7 \end{vmatrix}}{\begin{vmatrix} 0.6 & -0.4 \\ -0.5 & 0.7 \end{vmatrix}} = \frac{1.1}{0.22} = 5.$$

Problem 2

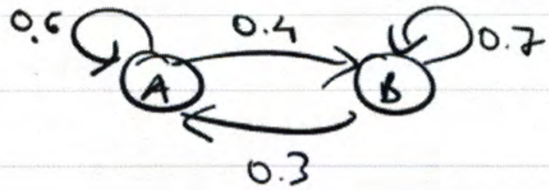
The average profit rate for a rental from location A is:

$$r_A = 0.6 \times 15 + 0.4 \times 25 = 19 \ \$/\text{day}$$

Similarly the average profit rate for a rental from location B is

$$r_B = 0.7 \times 15 + 0.3 \times 25 = 18 \ \$/\text{day}$$

The transition of any company vehicle among the locations A and B, between two consecutive rentals of the vehicle, is given by the following DT-MC:



Let π_x = the long-run proportion of rentals of the vehicle that take place at location x, $x = A, B$

Then, (π_A, π_B) is the stationary distribution of the above DT-MC, and therefore:

$$\pi_A = 0.6\pi_A + 0.3\pi_B \ ; \ \pi_A + \pi_B = 1 \Rightarrow \pi_A = 3/7 \ ; \ \pi_B = 4/7$$

Then, the average profit rate per rental day for the company is:

$$\pi_A r_A + \pi_B r_B = \frac{3}{7} \cdot 19 + \frac{4}{7} \cdot 18 = 18.43$$

Solution to Problem 3

Taken from S.M. Ross and E.A. Peköz "A Second Course in Probability"

Proof. Let d_k be the period of state k. Let n, m be such that $P_{ij}^{(n)} P_{ji}^{(m)} > 0$. Now, if $P_{ii}^{(r)} > 0$, then

$$P_{jj}^{(r+n+m)} \geq P_{ji}^{(m)} P_{ii}^{(r)} P_{ij}^{(n)} > 0$$

So d_j divides $r + n + m$. Moreover, because

$$P_{ii}^{(2r)} \geq P_{ii}^{(r)} P_{ii}^{(r)} > 0$$

the same argument shows that d_j also divides $2r + n + m$; therefore d_j divides $2r+n+m-(r+n+m) = r$. Because d_j divides r whenever $P_{ii}^{(r)} > 0$, it follows that d_j divides d_i. But the same argument can now be used to show that d_i divides d_j. Hence, $d_i = d_j$. ■

It follows from the preceding that all states of an irreducible Markov chain have the same period. If the period is 1, we say that the chain is **aperiodic**. It's easy to see that only aperiodic chains can have limiting probabilities.

Intimately linked to limiting probabilities are stationary probabilities. The probability vector $\pi_i, i \in S$ is said to be a **stationary probability vector** for the Markov chain if

$$\pi_j = \sum_i \pi_i P_{ij} \quad \text{for all} \quad j$$

$$\sum_j \pi_j = 1$$

Its name arises from the fact that if the X_0 is distributed according to a stationary probability vector $\{\pi_i\}$ then

$$P(X_1 = j) = \sum_i P(X_1 = j | X_0 = i)\pi_i = \sum_i \pi_i P_{ij} = \pi_j$$

and, by a simple induction argument,

$$P(X_n = j) = \sum_i P(X_n = j | X_{n-1} = i) P(X_{n-1} = i) = \sum_i \pi_i P_{ij} = \pi_j$$

Consequently, if we start the chain with a stationary probability vector then $X_n, X_{n+1}, \ldots$ has the same probability distribution for all n.

The following result will be needed later.

Problem 4

1) We prove this result by induction on k.

The truth of the base case $(k=1)$ is obvious from the definition of $G(A)$.

Next we assume that the considered statement holds for directed paths of length up to k, and we shall show that it must also hold for paths of length $k+1$.

First consider a nodal pair (i,j) and suppose that there exists a directed path in $G(A)$ that leads from i to j and has length $k+1$. Let q denote the node of $G(A)$ reached by this path in k steps. Then from the induction hypothesis we have that $a_{i,q}^{(k)} > 0$.
Also, $a_{q,j} > 0$ since (q,j) is an edge of the considered path. Hence, $a_{i,q}^{(k)} \cdot a_{q,j} > 0$. (1)
But $a_{i,j}^{(k+1)} = \sum_{e} a_{i,e}^{(k)} \cdot a_{e,j} \geq a_{i,q}^{(k)} \cdot a_{q,j}$ (2)
since $A \geq 0$.

Finally, from (1) and (2), $a_{i,j}^{(k+1)} > 0$, and since (i,j) was chosen arbitrarily, we have established the forward part of the considered statement.

For the reverse part, consider (i,j) with $a_{i,j}^{(k+1)} > 0$. Since $a_{i,j}^{(k+1)} = \sum_\ell a_{i,\ell}^{(k)} a_{\ell,j}$, there exists ℓ s.t. $a_{i,\ell}^{(k)} a_{\ell,j} > 0$, which further implies that $a_{i,\ell}^{(k)} > 0$ and $a_{\ell,j} > 0$. Then, the sought result follows from these two inequalities, the definition of $G(A)$, and the induction hypothesis.

2) It suffices to show the result for $k' = k+1$. For this case, we prove the result by contradiction.
Hence, suppose that $a_{i,j}^{(k+1)} = 0$ for some pair (i,j). Since $a_{i,j}^{(k+1)} = \sum_\ell a_{i,\ell}^{(k)} a_{\ell,j}$, the working hypothesis together with the facts that $A \geq 0$ and $A^k > 0$ further imply that $a_{\ell,j} = 0$, $\forall \ell$; i.e., the entire j-th column of matrix A is equal to $\emptyset$. But then, the definition of matrix multiplication implies that every matrix A^n, $n = 1, 2, \ldots$, has its j-th column equal to zero, which contradicts the fact that $A^k > 0$. —

Also, in the semantics of $G(A)$, the fact that $a_{l,j} = 0$, $\forall l$, implies that node j cannot be reached by any node (including itself), and the fact that $A^K > 0$ contradicts part (1) of this problem.

Problem 5

Let $s(t)$ be the column vector that collects the expert scores at iteration t, $t = 1, 2, \ldots$. Also set $W = [w_{ij}]$ $i, j = 1, \ldots, n$. Then W is a stochastic matrix with nonzero diagonal elements, and we can represent the voting process by

$$s(t+1) = W \cdot s(t), \quad t = 1, 2, \ldots$$

Also,

$$s(t) = W^{t-1} s(1), \quad t = 2, 3, \ldots$$

and

$$\lim_{t \to \infty} s(t) = \left(\lim_{t \to \infty} W^{t-1} \right) s(1)$$

Let

$$\lim_{t \to \infty} W^{t-1} = \tilde{W}$$

Since W is stochastic, it defines a finite-state DT-MC. If this MC is irreducible, then the non-zero elements in the principal diagonal of W imply that it is also aperiodic, and, from the corresponding developments presented in class, $\tilde{W}$ is also primitive. Hence, if π is the row vector denoting the stationary distribution of the corresponding MC, we have

$$\tilde{W} = \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}$$

and $\lim_{t \to \infty} s_i(t) = \pi^T \cdot s(1) \quad \forall i \in \{1, \ldots, n\}$.

Hence, the irreducibility of W attains, asymptotically, the desired consensus. Furthermore, the convergence to the eventually agreed score occurs exponentially fast, which is another desired attribute of the scoring process.

Next, we generalize further the above condition to the requirement that the DT-MC induced by W has a single closed communicating class. In this case, possibly with some renumbering of the experts, W can be written as follows:

$$W = \begin{bmatrix} W_T & W_{TC} \\ 0 & W_C \end{bmatrix}$$

where: W_T captures the transitions of the induced DT-MC within its transient states; W_C captures the transitions of the DT-MC within the closed communicating class; and W_{TC} captures the transitions from transient states to the closed communicating class.

Also, in this case,

$$\tilde{W} = \lim_{t\to\infty} W^t = \begin{bmatrix} \lim_{t\to\infty} W_T^t & \tilde{W}_{TC} \\ \emptyset & \lim_{t\to\infty} W_C^t \end{bmatrix}$$

Since W_T^t is substochastic, all of its eigenvalues have a modulus strictly less than 1.0, and working as in the case of stochastic matrices, we can show that $\lim_{t\to\infty} W_T^t = \emptyset$. This is consistent with the fact that transient states of DT-MCs have zero limiting probabilities.

On the other hand,

$$\lim_{t\to\infty} W_C^t \equiv \tilde{W}_C = \begin{bmatrix} \pi_c \\ \pi_c \\ -- \\ \pi_c \end{bmatrix}$$

where π_c is the stationary distribution of the irreducible DT-MC that is induced by W_c. We claim that each row of $\tilde{W}_{TC}$ is equal to π_c as well.

To see this, first notice that since $\tilde{W}_c$ is the limit of stochastic matrices, all of its rows must add up to 1.0. And since $\lim_{t \to \infty} W_T^t = \emptyset$, every row of $\tilde{W}_{TC}$ defines a distribution on the states that belong in the closed communicating class. To see that these distributions are equal to π_c for each row of W_{TC}, just notice that the underlying DT-MC defined by W will enter the closed communicating class w.p. 1, and afterwards it will operate according to the dynamics defined by W_c.

Finally, letting $s(1) = \begin{bmatrix} s_T(1) \\ s_c(1) \end{bmatrix}$,

from the above discussion we have:

$$\lim_{t \to \infty} s(t) = \begin{bmatrix} 0 & \pi_c \\ 0 & \pi_c \\ \hline 0 & \pi_c \end{bmatrix} \begin{bmatrix} s_T(1) \\ s_c(1) \end{bmatrix} = \begin{bmatrix} \pi_c \cdot s_c(1) \\ \pi_c \cdot s_c(1) \\ \hline \pi_c \cdot s_c(1) \end{bmatrix}$$

A more intuitive interpretation of the above results is obtained by noticing that in the considered operational context, $W_{i,j} > 0$ implies that expert i is influenced by expert j in her scoring decisions, and the value of the weight $W_{i,j}$ signifies the strength of this influence.

Also, paths in the underlying state transition diagram express indirect, delayed influence.

Even more interestingly, recurrence corresponds to some notion of persistence, while transience implies that the corresponding experts do not influence the persistent ones but are only influenced by them.

Finally, it is easy to see from all the above that, in general, there will be no consensus if the DT-MC induced by W has two or more closed communicating classes. –